https://doi.org/10.1016/j.mineng.2019.106178

A neural network approach for spatial variation assessment – a nepheline syenite case study

Camilo A. Mena Silva [a,*], Steinar L. Ellefmo [a], Roar Sandøy [b], Bjørn E. Sørensen [a] and Kurt Aasly [a]

[a] Department of Geoscience and Petroleum, Norwegian University of Science and Technology, NO-7491Trondheim, Norway

[b] Sibelco Nordic AS, N-1309 Sandvika, Norway

* Corresponding author. Tel.: +4740745343.
E-mail address: camilo.silva@ntnu.no

Abstract

The present geometallurgical study shows the application of a machine-learning methodology to the prediction of material properties from the Nabbaren nepheline syenite deposit in Norway. The approach used in this study created and tested a shallow neural network along with cluster analysis for the prediction of laboratory concentrate yield and modal mineralogy. The input is bulk chemistry data from the mining company open pit database. The methodology proposed unveils general trends in the deposit to a suitable operational scale for the open pit mine. The accuracy of the prediction models is acceptable, with one of the prediction models achieving a correlation coefficient 0.9. The application of a neural network approach showed a successful attempt in the prediction of concentrate yield and modal mineralogy in the Nabbaren nepheline syenite deposit. However, further investigations in terms of deposit internal variation and mineralogical studies are needed for utilising these prediction models, and further improve the modal mineralogy prediction model by better domaining and more representative distribution of samples used for modal mineralogy analyses.

Keywords: Geometallurgy, Neural network, Concentrate yield, Mineralogy, Modelling, Spatial Modelling, Industrial minerals, Nepheline syenite.

1. Introduction

Geometallurgy is a holistic discipline aiming to incorporate information from the entire mining value chain including data describing the characteristics of the raw material, intermediate and final products, and the processing involved. By incorporating all the information, it is possible to reduce operational risks and increase resource performance (Cropp et al., 2014; Dominy and O'Connor, 2016; Dunham et al., 2011; Lang et al., 2018; Lund and Lamberg, 2014; Walters, 2011). The data used in geometallurgy is quantitative and georeferenced. In this way it can be integrated into a 3D model and used for production schedule development (Hunt and Berry, 2017; Lamberg, 2011). The acquisition of quantitative data is expensive both in terms of costs and time. Thus, data prediction through mathematical modelling is not only a way to overcome these inconveniences, but also an interesting field of research in the mining industry (Berry et al., 2015; Lund et al., 2015; Rosa et al., 2014; Suazo et al., 2010). These types of prediction models can also be regarded as geometallurgical models, which are developed through a geometallurgical program (Lishchuk et al., 2015a; Lishchuk et al., 2015b) and visualized in a geometallurgical flowsheet (Lang et al., 2018).

To develop geometallurgical models it is necessary to quantify the different characteristic properties of the material of interest like bulk chemistry or modal mineralogy. For example, it is known that mineralogical information can be linked to mineral processing performance indicators, like concentrate yield and product quality (Brough et al., 2013; Lotter et al., 2018; Malvik, 2014; Petruk, 2000; Schapiro et al., 1981).

Mineralogical quantifications like modal mineralogy and laboratory tests of mineral processing performance indicators are key geometallurgical features that are costly and time-consuming (e.g. sample preparation and data acquisition), thus any improvement during data acquisition implies an economical benefit, and a possibility for real-time raw material or product quality control. In this regard, finding a way to estimate results of laboratory tests of mineral processing performance and modal mineralogy data is a time and cost-effective alternative to time-consuming laboratory tests and modal mineralogy analysis of large representative sample increment sets.

At industrial mineral operations, where product specifications are mainly based on chemical content, bulk chemical analyses and laboratory tests of processing performance are often more common than modal mineralogy analyses. To relate bulk chemistry to modal mineralogy estimations and results of laboratory tests of processing performance, would be beneficial. Modal mineralogy estimations based solely on chemistry are known as element-to-mineral conversion (EMC), and have been developed during the last four decades (Berry et al., 2011; Berry et al., 2015; Bryan et al., 1969; Hestnes and Sørensen, 2012; Johnson et al., 1985; Parian et al., 2015; Whiten, 2007; Yvon et al., 1990). Often these are based on the assumption of a linear relationship between mineral processing performance indicators and mineralogy (Berry et al., 2011; Mena Silva et al., 2018; Whiten, 2007). However, mineral processing performance indicators can be explained with mineralogical data, unfortunately it often presents additivity issues (Dunham and Vann, 2007; Lishchuk et al., 2019). Examples of mineralogical data explaining processing performance are mineralogical textures and associations (Johnson et al., 2007; Koch et al., 2019; Lund et al., 2015; Mwanga et al., 2015; Pérez-Barnuevo et al., 2018; Tøgersen et al., 2018). Usually, the relation between mineralogical data and mineral processing performance is non-linear relationship, thus a non-linear estimation technique, like an artificial neural network framework, could favourably be used.

Pattern recognition or machine-learning techniques, like artificial neural networks, have been used in identification, prediction and control of properties (Bishop, 2016; Demuth et al., 2014; Goodfellow et al., 2016; Masters, 1993) in a broad variety of fields like banking, data mining, climate and medicine. In the case of mineralogical and mineral processing studies, machine-learning techniques such as neural networks have been used in mineral classification based on X-ray data (Gallagher and Deacon, 2002; Koujelev and Lui, 2011; Rozel et al., 2014; Tsuji et al., 2010), mineral recognition in a frame of geometallurgy (Koch et al., 2019; Leroy and Pirard, 2019; Pérez-Barnuevo et al., 2018), in mineral exploration (Rigol-Sanchez et al., 2003), and to relate spatially processing performance with ore characteristics (Lishchuk et al., 2019; Rajabinasab and Asghari, 2018).

The present research aims to develop and test a neural network methodology, based on bulk chemistry measurements, for the prediction of two performance indicators: the mass recovered in the concentrate from the feed material in percentage from magnetic separation laboratory tests, which is defined as concentrate yield, and the modal mineralogy from reverse circulation drill-hole samples. The development of this methodology aims to use bulk chemistry measurements from reverse drill-hole samples and their corresponding measurements due to the number of available samples in the company database. This kind

of geometallurgical utilisation of this data set would represent the fastest and most cost-efficient way of obtaining information on mineralogy and concentrate yield in many industrial mineral operations. As a test case in this study an industrial mineral mine located in northern Norway and exploited because of the economical value associated to the nepheline and K-feldspar content in their raw material was chosen. The mine currently uses bulk chemistry and laboratory mineral processing concentrate yield, using laboratory dry magnetic separation tests, to define raw material quality as an input to production planning and quality control. The study focuses on a methodology to estimate laboratory processing concentrate yield based on three different data inputs; bulk chemical data from the company open pit database, a location class based on lithological map and estimated modal mineralogy.

2. Background

2.1. Geological framework

The material in this study was obtained from the Nabbaren deposit, a nepheline syenite deposit, located on the Island of Stjernøy, Northern Norway (Fig. 1).

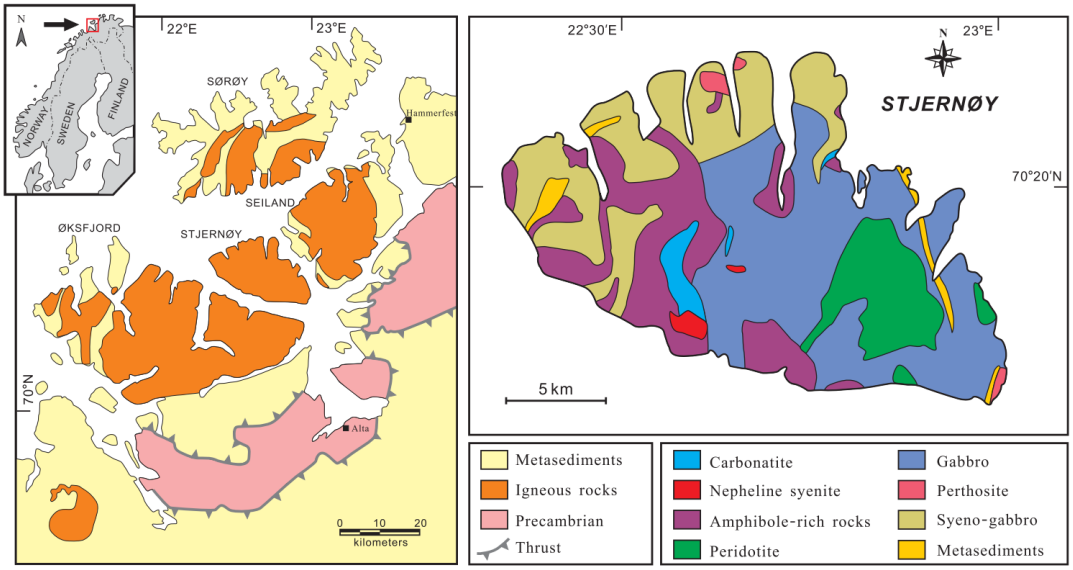

Fig. 1. *On the left: Geological interpretation of the Seiland igneous province, Northern Norway modified by Li (2013) after Robins and Gardner (1975). On the right: Simplified geological map of the Stjernøy island from Li (2013).*

The Nabbaren deposit is located within the Lillebukt alkaline complex in the Seiland province. The main rock types of the complex are hornblende clinopyroxenite, alkali syenite and carbonatite (Robins and Often, 1996). The main lithologies surrounding the deposit are nepheline-syenite gneiss to the northeast and amphibole-rich syenitic fenite to the southwest

(Fig. 2). The deposit dips steeply and is elongated in a northwest-southeast direction, with a length of 1700 metres along strike and a width of 300 metres (Geis, 1979).

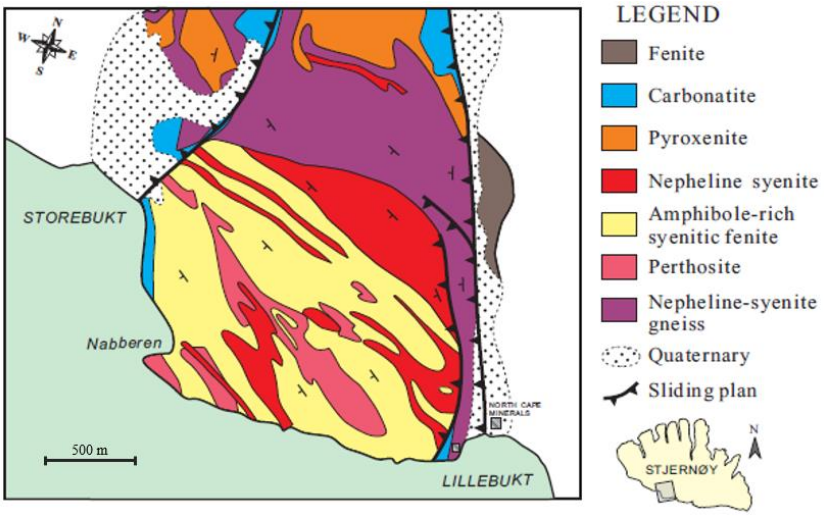

Fig. 2. *Detailed geological map of the Lillebukt alkaline complex on the Stjernøy Island from Li (2013).*

According to Geis (1979) the carbonatites and nepheline-rich gneisses are the main rock types on the northeast side of the Nabbaren deposit, whereas the south-western part is dominated by amphibole-rich lithologies.

The Nabbaren nepheline syenite is in contact with a nepheline-syenite gneiss that has biotite and hornblende as the main mafic minerals, see Fig. 2. The nepheline-syenite gneiss has been referred to as either a skarn gneiss high in calc-silicates (Heier, 1961) or as a carbonatite-nepheline syenite (Oosterom, 1963) because of its characteristic nepheline-albite and calcite rich bands (Heier, 1961).

The hornblende-rich syenitic fenite located in the southeast of the Nabbaren deposit, is rich in magnesium and iron. The amount of biotite in the syenitic fenite is associated to desilication, which decreased the plagioclase and K-feldspar content that additionally enriched the biotite and amphibole with potassium (Robins and Tysseland, 1979).

The two main minerals found in the Nabbaren deposit at Stjernøy are alkali feldspar (avg. 56 wt. %) and nepheline (avg. 34 wt. %). Common accessory minerals in the deposit are biotite, plagioclase (albite), calcite, magnetite, clinopyroxene, hornblende and titanite (Geis, 1979; Heier, 1961). Trace minerals have been identified and are found in different zones in the deposit (e.g. faults, mafic dike, etc.). The trace minerals include Al_2O_3-rich salite (diopside), Ca-rich amphibole, apatite, ilmenomagnetite, ilmenite in hornblende clinopyroxenites, diaspore, natrolite and thomsonite, as well as stronalsite-banalite (Li, 2013; Robins and Often, 1996). The mineral chemistries from the Nabbaren deposit have been quantified in numerous studies (Barth, 1963; Heier, 1962, 1964, 1966; Li, 2013; Mena Silva et al., 2018; Mjelde, 1983; Sørensen, 2016). Investigations on the mineral chemistry variation in the deposit have been documented for K-feldspar, nepheline and albite in terms of Sr and Ba contents, as shown in Fig. 3. This variation in Sr and Ba could be associated to a magmatic fractionation process and the geological association of carbonatites with the nepheline syenite (Heier, 1964; Heier and Taylor, 1964).

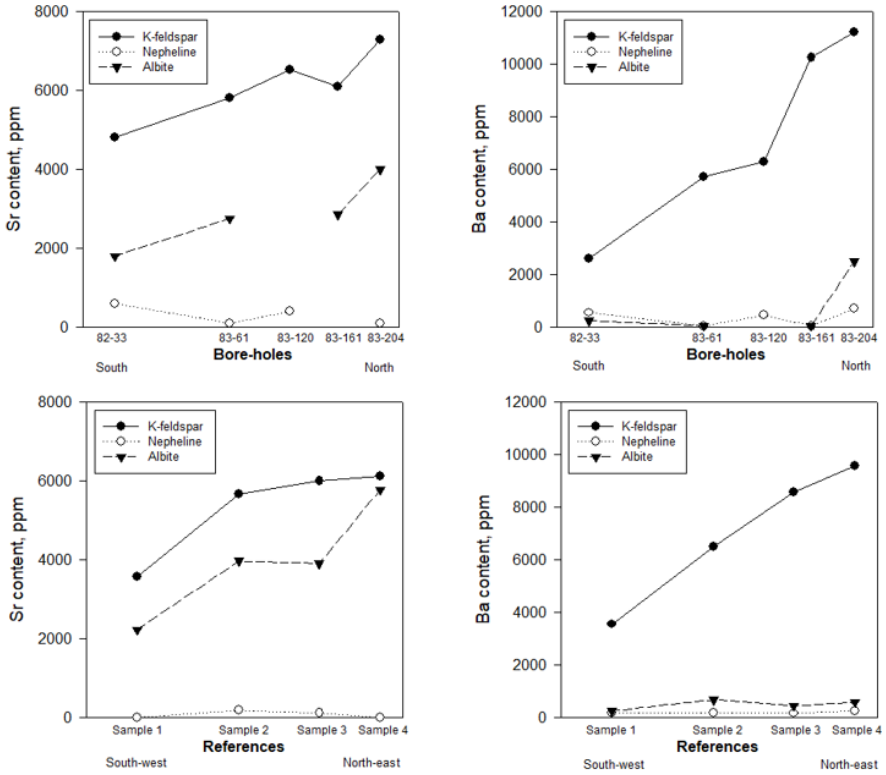

Fig. 3. *Mineral chemistry variation of K-feldspar, nepheline and albite within the Nabbaren deposit. Top-subplots from Mjelde (1983) based on Heier (1964). Bottom-subplots from Mena Silva et al. (2018).*

2.2. Mining and Mineral Processing

2.2.1. Mining and production planning

The deposit is mined both from an open pit and an underground mine (i.e. using an adapted open stoping as extraction method), from which roughly 80% and 20% feeds the processing plant, respectively. Both feed sources add up to a total of 550 kt of run of mine. Within the industrial mineral sector, production processes require strict quality control. Hence, strict mine planning and blending strategies are mandatory to avoid off-specification production or low product quality. According to Aasly and Ellefmo (2014), bad resource utilisation or worse, having an entire fully processed raw material batch discarded is a consequence of not reaching product specifications.

As part of mine planning and grade control procedures, a drill-in-advance campaign (DIA) is conducted before each mining season to plan feed material blending strategies and perform deposit model reconciliation for each production level in the open pit. The DIA campaign is performed using reverse-circulation (RC) drilling to extract a sample 12 metres long from a 10-by-10 metre grid. Each drill-hole extracts around 120 kilograms of drill cuttings. The drilling rig at the mine site does not collect material smaller than 75 µm. The drill cuttings from the reverse circulation drill-hole are split, and then characterised using concentrate yield from laboratory dry magnetic separation tests. Bulk chemistry quantifications are done on feed and concentrate (i.e. non-magnetic material) from the magnetic separation tests.

2.2.2. Mineral processing

Mineral processing at Stjernøy comprises comminution and mineral separation in a dry process to produce pure concentrates consisting mainly of nepheline and alkali feldspar (K-feldspar). A key feature making the deposit economically feasible is the inexpensive removal of Fe-bearing minerals in the rock, mainly associated to magnetite, biotite, hornblende and

clinopyroxene. The processing has multiple steps in comminution, particle size classification and magnetic separation. The mineral processing plant can be divided into (1) low-intensity magnetic separation for high magnetic susceptibility particles, and (2) high-intensity magnetic separation for low magnetic susceptibility particles. The plant annual production is more than 320 kt of mixed K-feldspar and nepheline concentrates. The concentrates are sold to a broad variety of industrial products like glass, ceramics and paint.

2.3. Routine data acquisition

2.3.1. Sample Preparation

In the laboratory magnetic separation internal procedure from the company uses a 2 kg subsample from each drill-hole sample. The magnetic separation test consists of seven magnetic separation steps with different settings similar to the company processing plant. The test performance of the samples is measured by its concentrate yield, a relative value calculated as the mass recovered in the concentrate from the feed material in a standardized magnetic separation laboratory test.

2.3.2. X-ray fluorescence (XRF)

The bulk chemistry data, obtained from the company, is done on pressed powder pills. The pills consist of 0.78 g of Licovax C and 6.50 g of representative sample. The samples are milled with a Fritsch Pulverisette 7 planetary mill. The pills are pressed using a Fluxana Vaneox manual press. The equipment used for the measurements was an ARL AdvantX Sequential spectrometer with OXSAS 2.2 software. The X-ray source used was a copper cathode.

The analytical routine is set up specifically for the nepheline syenite at Stjernøy and calibrated with a nepheline syenite reference provided by the British Ceramic Research and the Bureau of Analysed Samples, and labelled BCS-RM No 201a. The analyses are run with an electric current of 15 mA and a voltage of 40 kV. Acquisition time for each analysis is 242 seconds.

3. Experimental

3.1. Materials

The dataset used in this study corresponds to the DIA drill-hole bulk chemistry (XRF) and laboratory concentrate yield data collected between 2015 and 2017 campaigns. The total number of drill-holes is 1055, distributed in three production levels in the open pit. From those drill-holes 46 were selected at random for modal mineralogy measurements. The locations of these samples are displayed in Fig. 4. The company dataset used in this study contains analytical results of Fe_2O_3, TiO_2, CaO, K_2O, SiO_2, Al_2O_3, MgO, Na_2O, BaO and concentrate yield. As example, Fig. 5 shows the spatial distribution of MgO, Fe_2O_3, SiO_2 and CaO in the deposit. The summary statistics of the bulk chemistry and concentrate yield of the DIA drill-holes are displayed in Table 1.

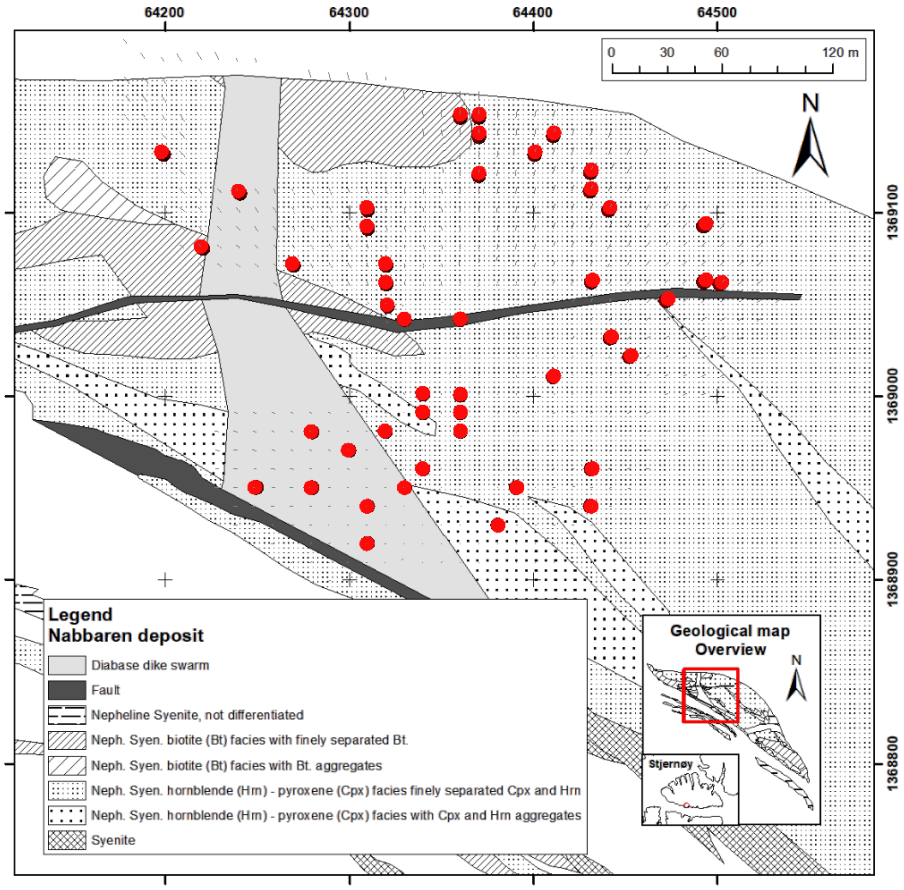

Fig. 4. *Sample position (in red) for modal mineralogy measurements in the detailed geological map of the Nabbaren nepheline syenite deposit. Geological map from Geis (1979) and updated by Sibelco Stjernøy with faults and the diabase dike swarm area.*

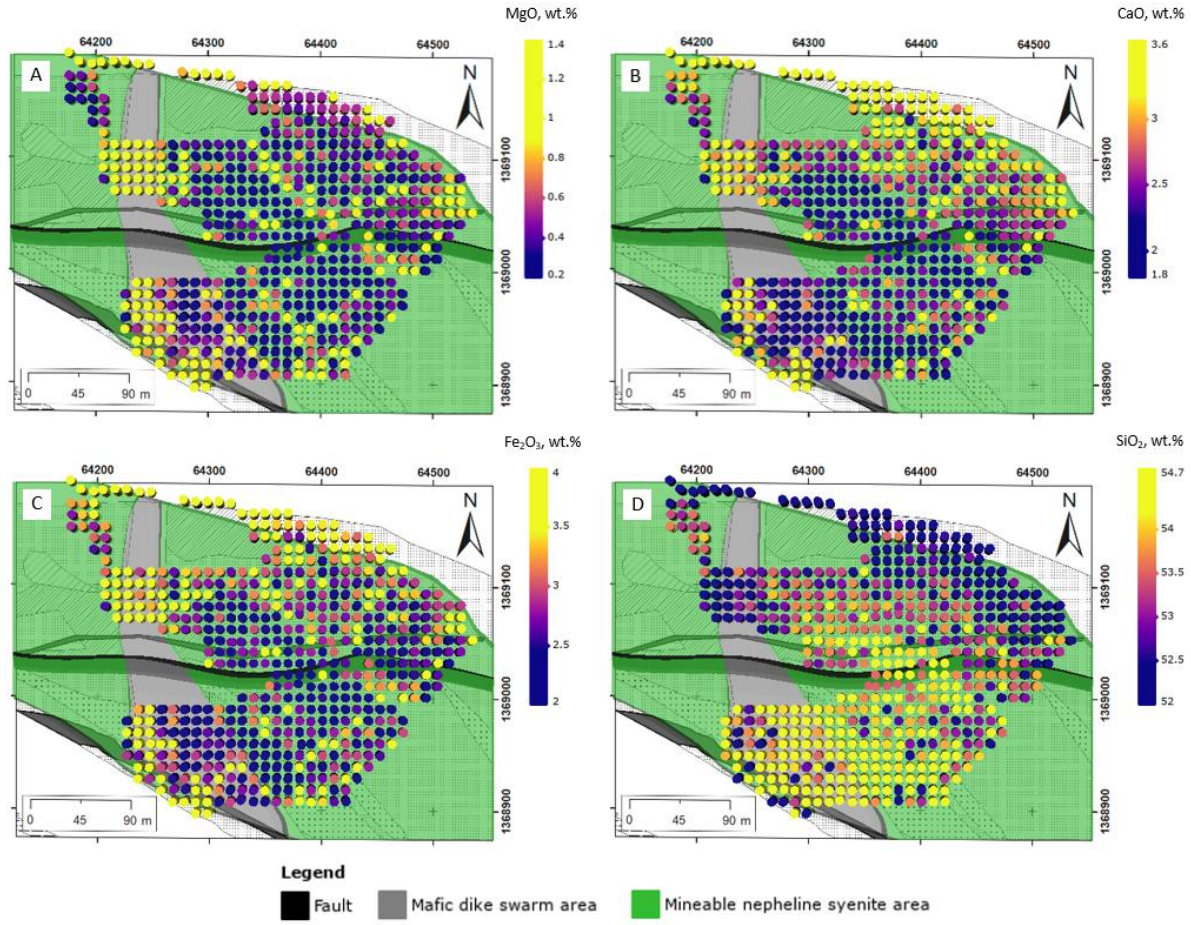

Fig. 5. *Open pit top view of the bulk chemical distribution of a) MgO wt.%, b) CaO wt.%., c) Fe_2O_3 wt.%, and d) SiO_2 wt.%. The shapes of the fault, mineable nepheline syenite area and mafic dike swarm areas vary with deposit depth. The legend in each plot is based on the histogram of every variable split into quartiles: the Q1 (lower quartile or 25th percentile of the data) is in the blue transition, the Q2 (median or 50th percentile) is in the purple transition, and the Q3 (upper quartile or 75th percentile) is in the yellow transition. The extreme values in the legend are a representation of the tails from each element histogram and correspond to the Q1 and Q3 values expanded by 0.5 wt.%.*

The deposit presents a significant spatial variability in the bulk chemistry of the samples in the north-east and south-west direction, see Fig. 5. Table 1 shows lithology independent summary statistics of the whole dataset of the DIA campaign between 2015 to the 2017.

Table 1 *Summary statistics of the 1055 points from the bulk chemical database for the 2015-2017 DIA campaigns.*

Oxide, wt.%	Mean	Std. dev.	Skewness
Al_2O_3	22.05	1.54	-2.98
BaO	0.34	0.08	-0.08
CaO	2.93	1.61	6.65
K_2O	8.09	0.88	-3.23
MgO	0.82	0.97	4.04
Na_2O	7.32	0.79	-1.87
SiO_2	53.20	1.91	-3.97
Fe_2O_3	3.16	1.25	2.88
TiO_2	0.70	0.32	3.21
Concentrate yield %	64.63	1.94	-2.21

In this study the mineable area was separated into the mineable nepheline syenite area and the mafic dike swarm area, and the data were assigned a location code, see Fig. 6. The mafic dike swarm area is a part of the nepheline syenite deposit where mafic dikes of variable thicknesses intersects the deposit (Ramsay and Sturt, 1970), resulting in an intimate mix of mafic dikes and nepheline syenite on a smaller scale than what can be separated by selective mining. From a mining perspective the area is mapped to signal an area where

product yield will be reduced due to contamination of the mafic dikes if the material is feed to the processing plant, since the dike material will end up in the tailings. The distribution of the location classes is shown in Fig. 6. The classification was based on the original lithological map of the deposit (Geis, 1979) and updated information from the mining operations. The summary statistics of the bulk chemistry and magnetic separation concentrate yield of the location classes are given in Table 2.

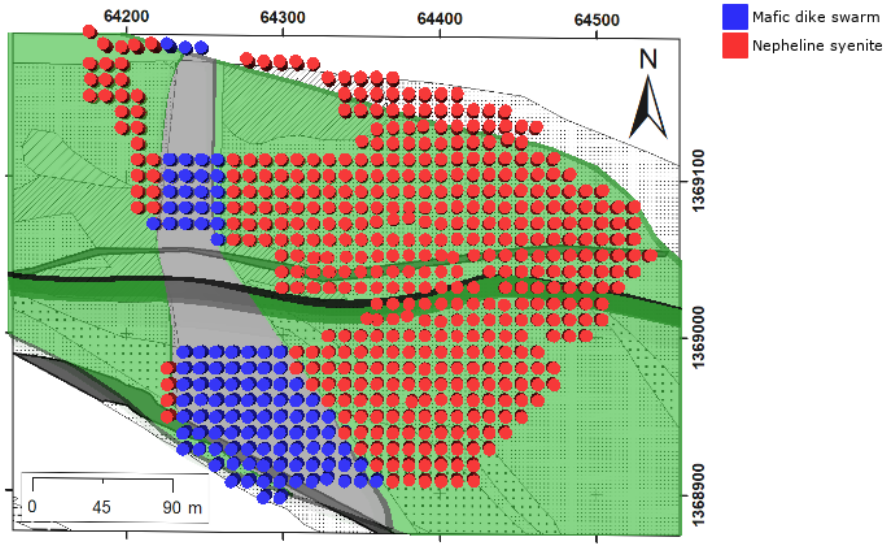

Fig. 6. *Open pit top view of the DIA drill-holes distribution according to location classes: nepheline syenite and mafic dike swarm area. The shapes of the fault (black), mafic dike swarm (grey) and mineable nepheline syenite (green) areas vary with deposit depth.*

The location classes show differences in bulk chemistry. The nepheline syenite class has higher content of Al_2O_3, K_2O, Na_2O, lower content of CaO, MgO and Fe_2O_3, and lower standard deviation variability than the mafic dike swarm class. Contamination of mafic dikes in the ore can be considered in terms of Fe_2O_3 and MgO, which is higher in the mafic dike swarm than in the nepheline syenite class. The data variability in the mafic dike swarm samples is higher than in samples from the nepheline syenite class. This variability is partially due to the presence of other contaminants like single mafic dikes or alteration zones in the nepheline syenite, not possible to distinguish on the map.

Table 2 *Summary statistics of the nepheline syenite and mafic dike swarm location classes from SIBELCO Stjernøy database for the 2015-2017 DIA campaigns.*

Location Class	**Nepheline syenite**			**Mafic dike swarm**		
Points	846			209		
Oxide, wt.%	Mean	Std. dev.	Skewness	Mean	Std. dev.	Skewness
Al_2O_3	22.21	1.31	-2.90	21.43	2.15	-2.39
BaO	0.36	0.08	-0.15	0.28	0.07	-0.01
CaO	2.87	1.39	7.25	3.19	2.29	5.01
K_2O	8.17	0.71	-2.93	7.78	1.34	-2.44
MgO	0.67	0.69	4.24	1.42	1.55	2.58
Na_2O	7.45	0.67	-1.88	6.80	1.00	-1.38
SiO_2	53.14	1.69	-4.16	53.46	2.63	-3.49
Fe_2O_3	3.05	1.03	2.88	3.64	1.83	2.07
TiO_2	0.67	0.27	2.58	0.80	0.44	3.22
Concentrate yield %	66.53	9.73	-2.32	56.90	16.24	-1.44

In order to implement the location class into the models, two extra columns were appended to the data for the nepheline syenite and mafic dike swarm classes. To indicate the class belonging to each point the corresponding column class had a 1 while the other class had a 0.

3.2. Methods

3.2.1. X-ray diffraction (XRD)

The modal mineralogy data was acquired with diffraction analyses and quantified with Rietveld refinement on samples consisting of approximately 1 g of micronized material. Each sample was micronized using a McCrone Micronizing Mill for 2.5 minutes after adding 10 mL of ethanol, producing an average P_{50} lower than 9.5 µm. The equipment used was a Bruker X-ray Diffractometer D8 Advance set to 40 kV and 40 mA. The X-ray source was a Cu anode generating K_{α} radiation with wavelengths $K_{\alpha1}$ = 1.5406 Å and $K_{\alpha2}$ = 1.54439 Å, and a $K_{\alpha1}/K_{\alpha2}$ ratio of 0.5.

The 2θ range in the diffractograms was from 3° to 65° with a step size of in 0.01°, a spinning speed of 60 rpm and 0.6 seconds counting time per increment. The total analysis time per XRD scan was 71 minutes. Rietveld modal mineralogy quantifications were performed with a standardized MS-DOS and TOPAS4.2 script routine. The routine used a fixed mineral list to quantify the modal mineralogy of each sample. The mineral list used in the routine comprised the 10 main minerals present in the deposit: nepheline, K-feldspar (orthoclase), albite, clinopyroxene (augite), hornblende, biotite, titanite, magnetite, pyrrhotite, calcite, and natrolite (tetra-natrolite). The values correspond to semi-quantifications obtained from mineral identification in the Bruker EVA® that were manually adjusted using the Rietveld fitting in the Bruker TOPAS4.2 software (Cheary and Coelho, 1992).

The mineral chemistry data obtained from mineral formula calculations based on EPMA analyses from Mena Silva et al. (2018) were included in the crystal structure models for Rietveld refinement. The crystal structures of the main minerals from the deposit were implemented instead of using standard values from the Bruker TOPAS4.2 software.

3.2.2. Data correlation

The relation between the measured modal mineralogy, concentrate yield and bulk chemistry data was calculated using three correlation coefficients: Pearson, Kendall and Spearman. These coefficients return values between -1 and 1. Coefficient values close to -1 or 1 indicates a strong relationship between parameters, whereas a coefficient value closer to 0 indicates a weak relationship. The sign of the coefficient value indicates whether the relationship is negative or positive. Pearson's coefficient, also known as linear correlation coefficient, defines the level of association between continuous variables assuming normal distribution. Unfortunately, Pearson's coefficient is seriously affected by the presence of outliers (Croux and Dehon, 2010). In this regard, Kendall and Spearman coefficients, known as rank correlation coefficients, have no dependency on the distribution of the measurements and estimate the level of association between ranked variables. The difference between Kendall and Spearman is that Kendall uses a relative order between ranks (i.e. qualitative), whereas Spearman is a linear correlation coefficient of the ranks (i.e. quantitative). Rank correlations are less sensitive to outliers in the dataset than linear correlations, more robust and efficient than linear correlations like Pearson's (Croux and Dehon, 2010; Press et al., 2007). The Matlab-function used to calculate these coefficients is *corr* and it includes options to specify the correlation type. All information on the functions can be found in Mathworks (2018a), and further information can be found in Press et al. (2007).

3.2.3. Data handling

The prediction models in this study were created following different workflows according to their objective. The work flows are shown in Fig. 7 for modal mineralogy modelling, and in Fig. 8 for concentrate yield modelling. Further information on the main steps in the work flow is in the subsequent sections.

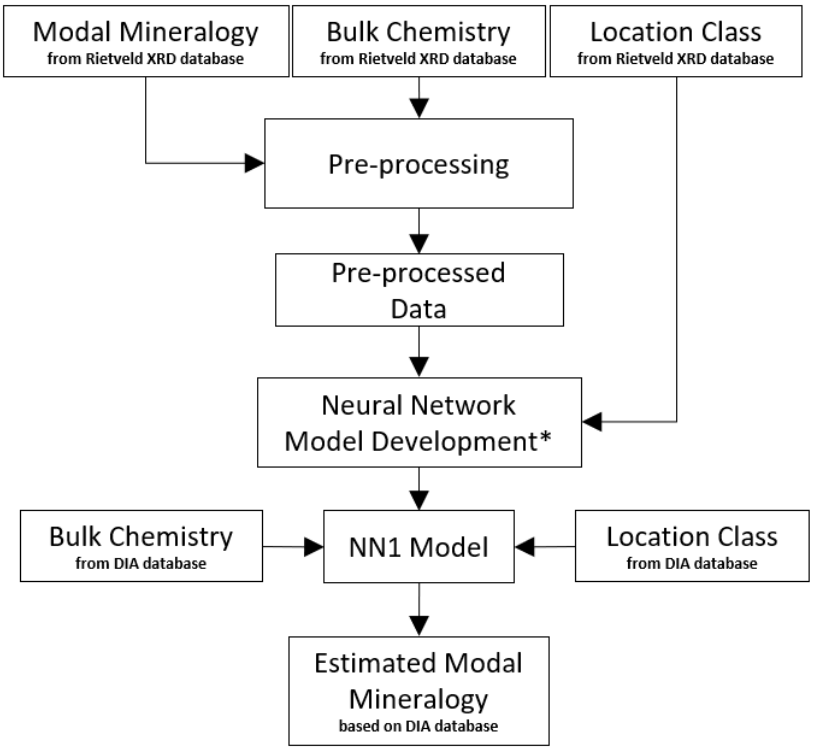

Fig. 7. *Data handling workflow for modal mineralogy modelling from Rietveld quantifications. * Processing (model definition) including number of neurons definition and network training.*

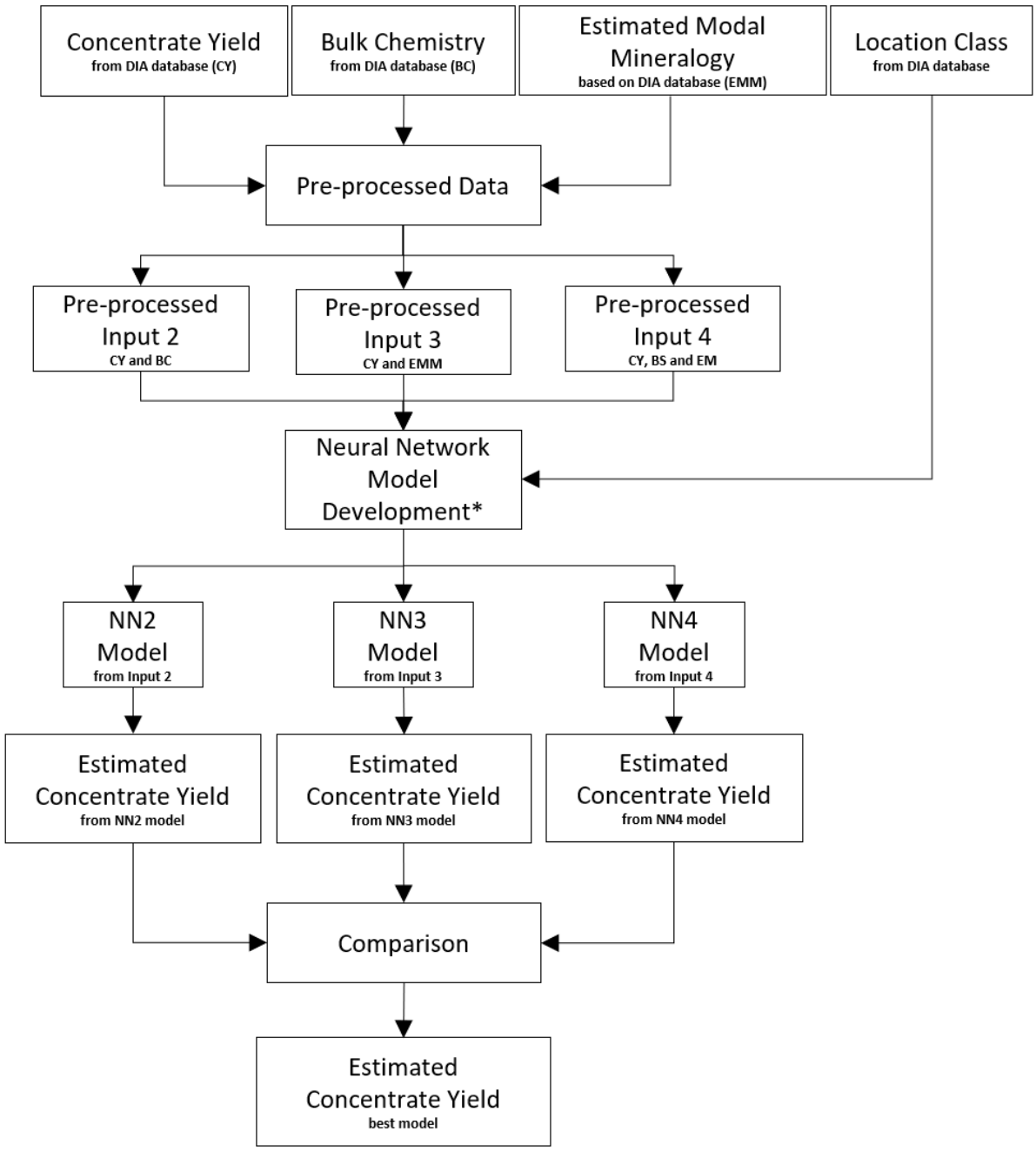

Fig. 8. *Data handling workflow for concentrate yield modelling from the DIA database. * Processing (model definition) including number of neurons definition and network training.*

3.2.3.1. Pre-processing (standardization and definition of subsets)

Before performing any estimation, data from both the DIA drilling campaign and Rietveld XRD mineralogy sets were standardized. The standardization procedure forces the input data to have zero mean and unit variance by subtracting the mean ($\mu_{i\ dataset}$) from the measured feature (x_i) in the dataset and dividing it by the standard deviation ($\sigma_{i\ dataset}$) of the measured feature in the dataset (Goodfellow et al., 2016), as shown in Eq. 1. The application of a standardization procedure is recommended in general for this kind of exercise because it makes the variables equally important, facilitates the interpretation of network weights (Goodfellow et al., 2016; Masters, 1993), and eases the learning process of

the network (Ioffe and Szegedy, 2015; Lecun et al., 1998). Nonetheless, it is not strictly required as shown by Karunanithi et al. (1994).

$$x_{standard\ i} = \frac{x_i - \mu_{i\ dataset}}{\sigma_{i\ dataset}} \quad (1)$$

The standardized datasets were split into three subsets: training, testing and validation (Maier et al., 2000). It is standard procedure to divide the dataset into training and validation sets, where the training data are used to create the model and the validation set to evaluate its performance. Subsequently, the validation set can be sub-divided into testing and validation subsets for model performance cross-validation. In this case the validation subset is used to monitor the model during training, whereas the testing subset is used after the model has been created to evaluate the model performance on untrained data (Maier et al., 2000).

A cluster analysis procedure was applied prior modelling to represent accurately the original datasets in each subset. The procedure used a hierarchical clustering using the inner squared distance between group pairs based on the Euclidean distance between points. This grouping technique is based on within cluster similarity (Ward, 1963) by merging groups according to any function reflecting the difference between groups (Ward, 1963), in this case the incremental sum of squares. The sum of squares is related to the distance between the components within the cluster and the cluster's centroid, which increases as a consequence of group merging. The methodology used the Matlab-function *linkage* with *ward* as the algorithm for computing the distance between clusters.

To define the number of clusters a cluster-tree or dendrogram (Fig. 9) was inspected visually. The number of clusters is defined according to the inconsistency (i.e. the heights or vertical lines) between cluster pairs (i.e. the links or horizontal lines), the higher the inconsistency the further apart the pair is from the rest. The number of clusters defined from the dendrogram visual inspection was afterwards used as the maximum number of clusters to set in the Matlab function *cluster* (i.e. using *maxclust* criterion). Once the number of clusters was defined, sample points were selected randomly from each cluster and allocated into each of the three subsets. The training, testing and validation subsets represented 50%, 25%, and 25% of the original dataset points, respectively.

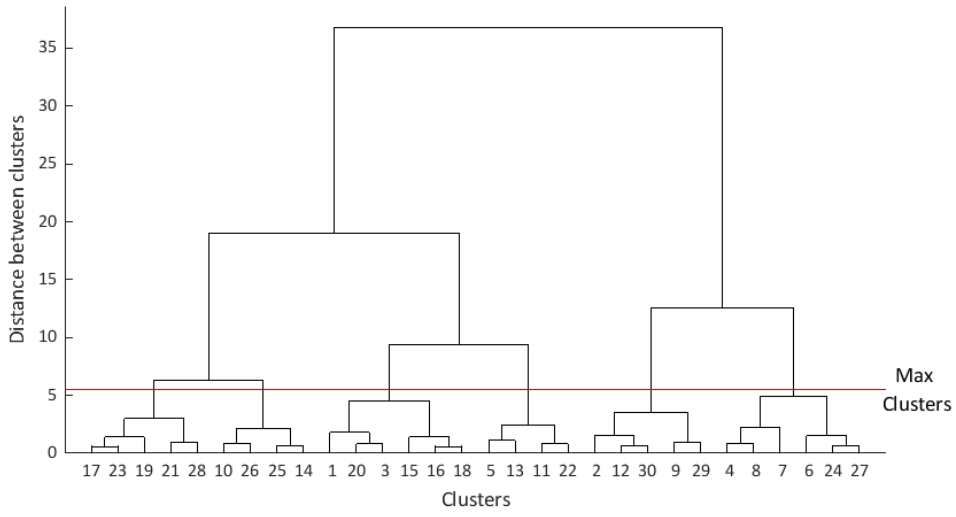

Fig. 9. *Dendrogram for hierarchical clustering of concentrate yield data. Max-clusters line indicates the number of clusters selected.*

In addition to the bulk chemical and modal mineralogy data, location classes (Fig. 6) and an *unknown* variable were added to the original data. The *unknown* variable, included as part of the bulk chemical data, represent all the elements *not* measured by the XRF analyses and the possible quantification errors, and was calculated by subtracting the sum of all the measured elements from the total, see Eq. 2.

$$unknown = 100\% - \sum measured\ element \quad (2)$$

3.2.3.2. Processing (model definition)

To estimate concentrate yield and modal mineralogy a neural network methodology was selected because they are considered as a non-linear approach to estimate the relationship between datasets that perform better than linear approaches. An example of these differences has been shown by Mekanik et al. (2013).

Fig. 10 illustrates the structure of a neural network where the *input layer* is the first layer to the network and consists of input features or input parameters (black-filled circles) such as TiO_2, Fe_2O_3, etc. feeding the network structure. The *hidden layers* are the intermediate layers in the network and they consist of activation units or neurons (white-filled circles) that are functions relating the parameters or neurons from the previous to the next layer. The number of hidden layers and neurons in the hidden layers depend on the desired network structure. The *output layer* is the last layer in the network and consists of the features or parameters (grey-filled circles) such as concentrate yield, modal mineralogy, etc. predicted by the network. For more background information on neural network design see Demuth et al. (2014) or Goodfellow et al. (2016).

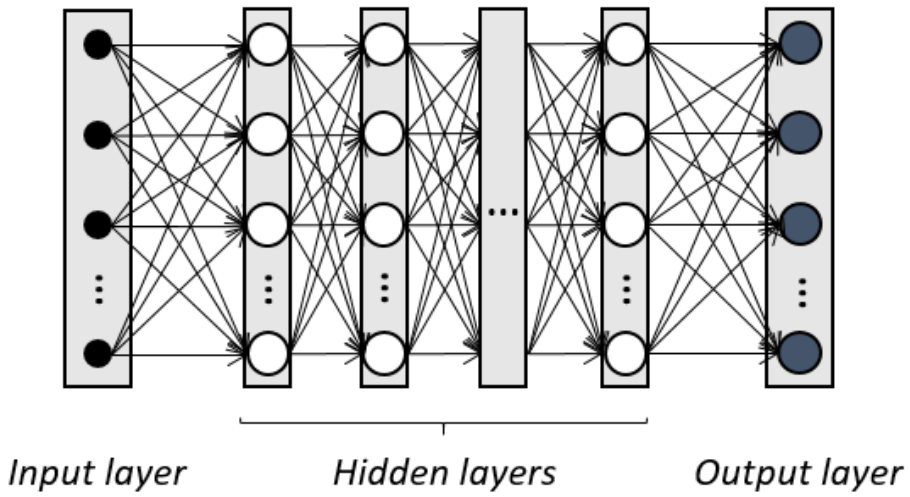

Fig. 10. *Structure scheme of a neural network.*

The model used in this study consisted of a feedforward neural network that was built with the *fitnet* Matlab-function. The *fitnet* function creates a network with a defined number of hidden sizes (or neurons) and a training function for its optimization. In this study *two* hidden sizes and the *trainlm* function were selected. The *trainlm* used the Levenberg-Marquardt algorithm (Demuth et al., 2014; Marquardt, 1963) as its function for optimization. Additionally, the algorithm was selected in accordance to recommendations from exercises of similar characteristics (Parian et al., 2015) and because of the efficiency of the Levenberg-Marquardt algorithm that works as a two-in-one algorithm comprising the steepest gradient decent and Newton-Gaussian methods (Mathworks, 2019). Once the network is defined, it requires training to generalize (or predict) the desired output from the

given inputs. The training was done using the *train* Matlab-function. Complex network structures (i.e. many hidden layers and neurons) tend to perform well (Demuth et al., 2014; Maier et al., 2000), but neural networks with simpler structures tend to predict any output. This is the reason for the simple network structure chosen in this study.

The number of neurons used in the hidden layer of each model were determined by evaluating the network performance. The number of neurons were selected by a visual assessment of the average performance based on the best compromise between training and test performances by number of neurons using the same sub-sets to avoid data-leakage, as shown in Fig. 11. The performance of the models is measured in terms of mean square error (i.e. *mse* as the performance function of the network), see Eq.3. Where $p_{measured}$ and $p_{estimated}$ are the measured and estimated parameters respectively, R is the number of points used and ε is the *mse*. The network training process has multiple stopping criteria to avoid an excessive time consumption or unneeded optimized value. For further information regarding these criteria refer to Mathworks (2018b).

$$\varepsilon = \frac{\sum_R (\boldsymbol{p}_{measured} - \boldsymbol{p}_{estimated})^2}{R} \quad (3)$$

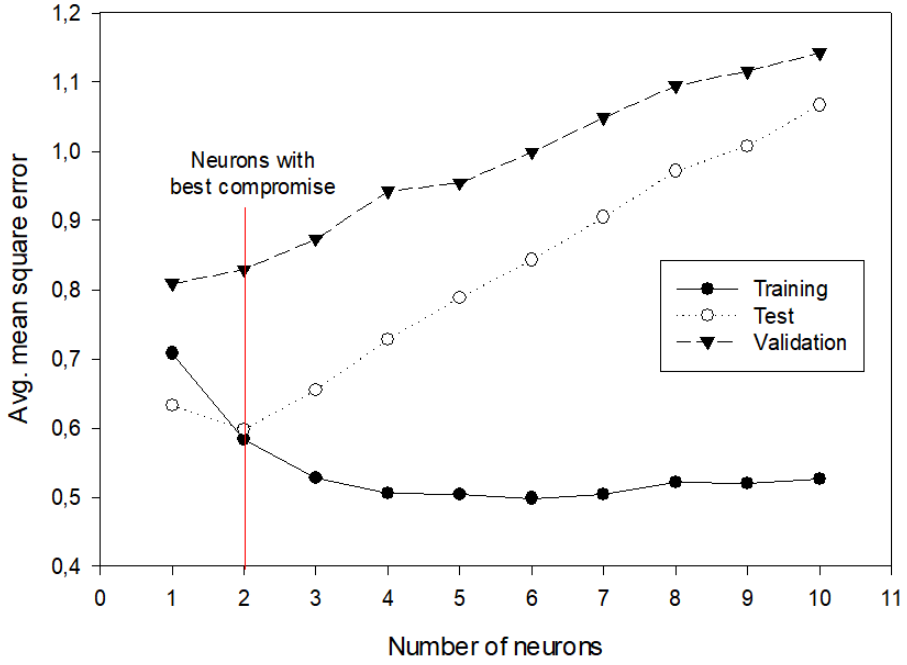

Fig. 11. *Selection of neurons for the hidden layer based on the best compromise between training and test average performance.*

If the prediction models estimated any negative values this was considered a flaw. When a negative value was estimated by the prediction model the whole input and output point was tagged as negative due to the model failure in estimating mineral quantities, and to have an idea of the location and range of these points.

3.2.3.2.1. Modal mineralogy estimations

The modal mineralogy estimations were based on 46 randomly selected samples from the deposit and the locations shown in Fig. 4. The modal mineralogy neural network model (or NN1) was based on bulk chemistry (CaO, TiO_2, Fe_2O_3, *unknown*, etc.),the location classes and the modal mineralogy measurements from Rietveld XRD refinement.

3.2.3.2.2. Concentrate yield estimations and comparison

Three estimation models were developed for laboratory concentrate yield, using three different input sets. The first model (NN2) used the location class and bulk chemistry. The second model (NN3) used the location class and estimated mineralogy (nepheline, biotite, titanite, etc.). The last model (NN4) used location class, bulk chemistry, and estimated mineralogy as input. A summary of the input data used in each concentrate yield estimation model is displayed in Table 3.

Table 3 *Input data for the concentrate yield estimation models.*

Model	***Input***	***Data***
NN 2	2	Bulk chemistry, location class
NN 3	3	Estimated mineralogy, location class
NN 4	4	Bulk chemistry, estimated mineralogy, location class

After creating and comparing the three concentrate yield models, the model with the lowest *mse* and fewest number of negative estimates was defined as the best model. The best model was further evaluated by checking the deviation between estimations and measured in ranges to see the samples by group. The ranges will group the estimated points according to their similarity towards the measured values, hence having an idea of how good the estimation was. According to the deviations four *indicator* groups were defined as displayed in Table 4. The reason for the creation of these groups is because the input data has a relationship with the performance of the model. The location of the input data in the deposit might indicate the reason of the estimation variation.

Table 4 *Indicator groups to evaluate estimation performance.*

Group	**Variability**
Best model	Difference < 2σ
+2sd indicator	2σ < Difference < 3σ
+3sd indicator	3σ < Difference < 4σ
+4sd indicator	4σ < Difference

4. Results

4.1. Mineralogy

Summary statistics of the 46 samples Rietveld XRD measurements and for the same dataset split into the two location classes are given in Table 5. Mineralogy by location class shows differences that otherwise would be masked, like the higher concentration of feldspar and albite in the mafic dike and the higher concentration of nepheline and titanite in the nepheline syenite. Regardless of the location class the modal mineralogy shows high variability with high standard deviations and skewness deviating from zero.

Table 5 *Summary statistics of the full XRD Rietveld dataset, and the same dataset split by location class. Nph = nepheline, Fsp = K-feldspar, Ab = albite, Cpx = clinopyroxene, Hbl = hornblende, Bt = biotite, Ttn = titanite, Mgt = magnetite, Po = pyrrhotite, Cal = calcite, Ntr = natrolite, based on Whitney and Evans (2010) nomenclature.*

	Rietveld XRD dataset			***Location class***					
				Nepheline syenite			**Mafic dike swarm**		
Points	46			38			8		
Mineral, wt.%	Mean	Std. dev.	Skewness	Mean	Std. dev.	Skewness	Mean	Std. dev.	Skewness
Nph	30.68	3.42	-1.94	31.11	3.46	-2.36	28.63	2.52	-2.11
Fsp	41.81	3.40	-0.08	41.71	3.41	0.10	42.32	3.51	-1.22
Ab	12.89	2.57	0.84	12.30	2.32	1.28	15.73	1.69	2.48
Cpx	4.87	2.04	0.27	5.16	2.05	0.13	3.51	1.38	-0.28
Hbl	2.99	3.11	2.29	2.96	3.09	2.43	3.11	3.41	2.16
Bt	1.91	1.11	1.39	1.84	1.09	1.60	2.26	1.26	0.81
Ttn	0.46	0.27	0.72	0.48	0.29	0.60	0.35	0.16	-0.48

Mgt	0.98	0.22	0.64	1.01	0.23	0.37	0.84	0.07	0.17
Po	0.30	0.10	-0.39	0.30	0.10	-0.33	0.29	0.11	-0.72
Cal	1.86	0.56	0.58	1.88	0.60	0.54	1.78	0.35	-0.48
Ntr	1.24	0.82	2.04	1.25	0.86	2.13	1.18	0.68	0.87

4.2. Data correlation

The correlation coefficients calculated between measured modal mineralogy, bulk chemistry and concentrate yield may indicate the type of relation between these parameters, and thus the type of methodology suitable to develop a prediction model. The values of the Pearson, Kendall and Spearman correlations are displayed in Table 6, Table 7 and Table 8, respectively.

Table 6 *Pearson correlation coefficient between mineralogy, bulk chemistry (including unknown) and concentrate yield. Nph = nepheline, Fsp = K-feldspar, Ab = albite, Cpx = clinopyroxene, Hbl = hornblende, Bt = biotite, Ttn = titanite, Mgt = magnetite, Po = pyrrhotite, Cal = calcite, Ntr = natrolite, based on Whitney and Evans (2010) nomenclature.*

Component, wt.%	***Al_2O_3***	***BaO***	***CaO***	***K_2O***	***MgO***	***Na_2O***	***SiO_2***	***Fe_2O_3***	***TiO_2***	***Unknown***	**Concentrate yield, %**
Nph	0.31	0.34	-0.15	0.25	-0.73	0.73	-0.04	-0.50	-0.40	0.22	0.51
Fsp	0.67	-0.35	-0.59	0.74	-0.45	-0.06	0.64	-0.58	-0.60	-0.64	0.52
Ab	-0.31	-0.45	-0.08	-0.14	0.40	-0.08	0.10	0.16	0.08	0.01	-0.21
Cpx	-0.24	0.60	0.44	-0.27	-0.11	-0.10	-0.22	0.20	0.25	0.23	-0.16
Hbl	-0.53	-0.05	0.36	-0.58	0.77	-0.36	-0.42	0.68	0.65	0.23	-0.59
Bt	-0.44	-0.32	0.24	-0.39	0.83	-0.42	-0.27	0.54	0.40	0.09	-0.49
Ttn	-0.35	0.79	0.58	-0.48	-0.14	0.08	-0.53	0.29	0.49	0.58	-0.39
Mgt	0.22	0.09	-0.11	0.23	-0.15	-0.08	0.17	-0.19	-0.27	-0.17	0.15
Po	-0.03	0.41	0.22	-0.01	-0.42	0.08	-0.04	-0.03	0.06	0.13	0.04
Cal	-0.37	0.40	0.52	-0.32	0.05	-0.04	-0.41	0.26	0.35	0.43	-0.28
Ntr	0.39	-0.02	-0.04	0.05	0.05	-0.38	0.11	-0.06	-0.05	-0.31	-0.06
Concentrate yield, %	0.63	-0.38	-0.66	0.75	-0.65	0.52	0.50	-0.74	-0.76	-0.37	1.00

Table 7 *Kendall correlation coefficient between mineralogy, bulk chemistry (including unknown) and concentrate yield.*

Component, wt.%	***Al_2O_3***	***BaO***	***CaO***	***K_2O***	***MgO***	***Na_2O***	***SiO_2***	***Fe_2O_3***	***TiO_2***	***Unknown***	***Concentrate yield, %***
Nph	0.18	0.22	-0.08	0.16	-0.42	0.37	-0.02	-0.21	-0.20	0.10	0.26
Fsp	0.51	-0.31	-0.56	0.65	-0.51	0.20	0.54	-0.59	-0.57	-0.50	0.45
Ab	-0.18	-0.31	-0.08	-0.06	0.21	-0.22	0.09	0.05	0.01	-0.05	-0.10
Cpx	-0.11	0.43	0.24	-0.17	-0.04	0.08	-0.15	0.10	0.16	0.18	-0.11
Hbl	-0.38	0.03	0.34	-0.48	0.62	-0.25	-0.35	0.45	0.45	0.23	-0.35
Bt	-0.19	-0.25	0.16	-0.21	0.50	-0.22	-0.12	0.28	0.21	0.01	-0.25
Ttn	-0.20	0.56	0.38	-0.39	0.10	0.07	-0.42	0.28	0.39	0.42	-0.34
Mgt	0.13	0.10	-0.04	0.14	-0.11	0.07	0.10	-0.09	-0.13	-0.08	0.09
Po	-0.14	0.30	0.15	-0.08	-0.03	0.10	-0.07	0.10	0.17	0.12	-0.06
Cal	-0.33	0.28	0.36	-0.25	0.17	-0.15	-0.29	0.26	0.29	0.32	-0.23
Ntr	0.26	0.02	-0.08	0.10	-0.07	0.04	0.12	-0.13	-0.06	-0.17	0.06
Concentrate yield, %	0.42	-0.29	-0.49	0.56	-0.50	0.32	0.37	-0.55	-0.57	-0.31	1.00

Table 8 *Spearman correlation coefficient between mineralogy, bulk chemistry (including unknown) and concentrate yield.*

Component, wt.%	***Al_2O_3***	***BaO***	***CaO***	***K_2O***	***MgO***	***Na_2O***	***SiO_2***	***Fe_2O_3***	***TiO_2***	***Unknown***	**Concentrate yield, %**

Nph	0.29	0.36	-0.11	0.24	-0.55	0.52	-0.03	-0.30	-0.29	0.15	0.35
Fsp	0.70	-0.43	-0.73	0.82	-0.67	0.30	0.70	-0.71	-0.71	-0.67	0.61
Ab	-0.25	-0.46	-0.12	-0.06	0.27	-0.29	0.15	0.07	0.00	-0.08	-0.14
Cpx	-0.13	0.56	0.36	-0.23	-0.04	0.11	-0.20	0.16	0.23	0.26	-0.12
Hbl	-0.54	0.07	0.49	-0.63	0.81	-0.37	-0.52	0.62	0.63	0.35	-0.49
Bt	-0.26	-0.35	0.18	-0.24	0.66	-0.34	-0.17	0.37	0.27	0.02	-0.34
Ttn	-0.32	0.75	0.55	-0.53	0.16	0.10	-0.56	0.38	0.53	0.57	-0.44
Mgt	0.18	0.15	-0.05	0.19	-0.16	0.12	0.11	-0.11	-0.16	-0.11	0.13
Po	-0.16	0.41	0.25	-0.10	-0.06	0.14	-0.10	0.14	0.22	0.19	-0.09
Cal	-0.44	0.42	0.53	-0.35	0.25	-0.22	-0.43	0.37	0.39	0.49	-0.31
Ntr	0.37	0.04	-0.14	0.17	-0.14	0.04	0.17	-0.20	-0.12	-0.25	0.09
Concentrate yield, %	0.58	-0.40	-0.69	0.77	-0.68	0.43	0.53	-0.74	-0.78	-0.43	1.00

Strong relationships are considered for any coefficient above 0.5 (or bellow -0.5). The strongest correlation values are observed between MgO with nepheline (negative), hornblende (positive), biotite (positive) and concentrate yield (negative); K_2O with K-feldspar (positive) and yield (positive); Na_2O with nepheline (positive); BaO with titanite (positive); TiO_2 and Fe_2O_3 with concentrate yield (negative). Based on data from Table 1, Table 2 and Table 5 and the distribution describing in them it would realistic to assume the suitability of Spearman and Kendall over Pearson's coefficients. This would indicate that the relationship between measured mineralogy, bulk chemistry and concentrate yield is mainly non-linear.

4.3. Subset definition

Cluster analysis was applied to the whole concentrate yield data from the company database, whereas to the Rietveld XRD data a cluster analysis was applied by location class (nepheline syenite and mafic dike swarm) because of the marked differences between one class to the other. From the cluster analyses six clusters were used for the concentrate yield, four clusters for the nepheline syenite class and three clusters for the mafic dike swarm class. An example of the similarity between subsets and original set distributions is shown in Fig. 12 and Table 9.

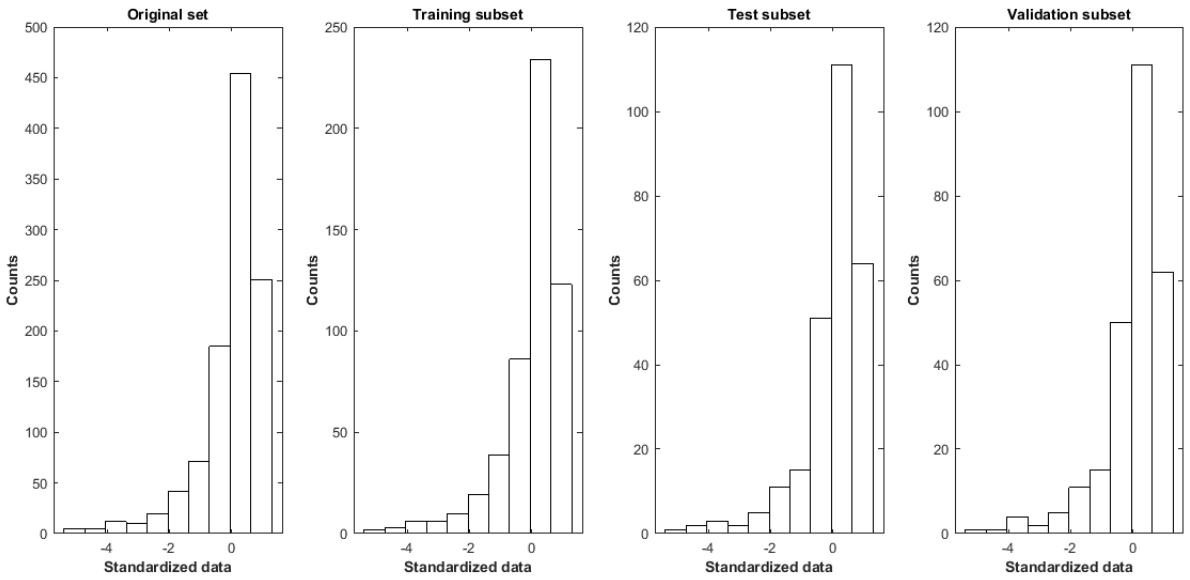

Fig. 12. *Concentrate yield histogram with the distribution of standardized points (counts) in the original set and its allocated subsets after cluster analysis.*

Table 9 *Standardized concentrate yield summary statistics of the original set and its allocated subsets after cluster analysis.*

	Original set	Training subset	Test subset	Validation subset
Samples	1055	528	265	262
Mean	2.8E-15	-1.4E-15	-2.4E-15	5.8E-17
Standard Deviation	0.03	0.04	0.06	0.06
Skewness	-2.21	-2.20	-2.26	-2.21
Minimum	-5.31	-4.96	-5.36	-5.22
Maximum	1.27	1.27	1.24	1.21

4.4. Model structures and performances

The number of neurons showing the best compromise between test and training performance, the final *mse* based on the test subset (i.e. untrained data) and the number of negative estimations from each model are displayed in Table 10.

Table 10 *Number of neurons, model performance and estimation of negatives. *Based on the average from each mineral.*

Model	***Input***	***Output***	***Neurons***	***mse***	***Negative estimations from DIA database***
NN 1	Bulk chemistry, location class	Mineralogy	2	0.66*	11
NN 2	Bulk chemistry, location class	Concentrate yield	2	0.16	0
NN 3	Estimated mineralogy, location class	Concentrate yield	4	0.20	0
NN 4	Bulk chemistry, estimated mineralogy, location class	Concentrate yield	5	0.16	2

4.5. Modal mineralogical estimations

The correlation between measured and predicted mineralogical data can give an idea of the model performance. High positive correlation indicates good model performance. Differences between estimated and measured nepheline, hornblende, biotite, clinopyroxene, K-feldspar and albite are shown in (Fig. 13).

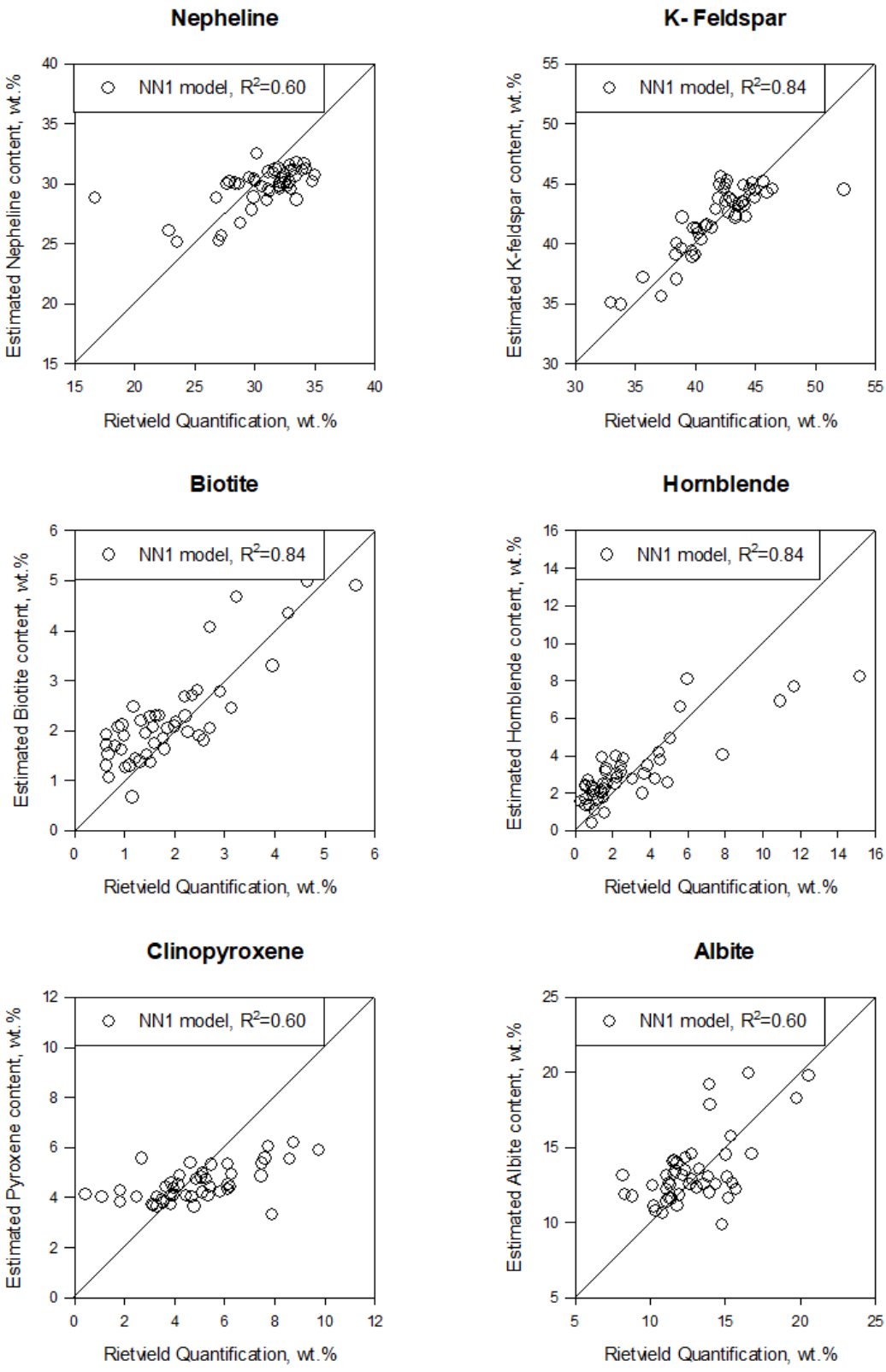

Fig. 13. *The Pearson correlation coefficient between the Rietveld refined XRD measurements and their respective estimated quantities from the NN1 model. Nepheline, hornblende, biotite, clinopyroxene, K-feldspar and albite are shown due to quantify range they cover which represent the performance of all the minerals used in this study*

The NN1 prediction model was applied to the full bulk chemical dataset of the company database to predict the mineralogy of each sample point. When applying the NN1 model to the whole bulk chemical dataset the modal mineralogy of 11 sample points were negatively

predicted on one or several bulk chemical parameters. The sample points with negative estimated values are mainly located in areas further apart from the samples collected for Rietveld XRD quantification or along the northern boundary of the deposit and in the southern part of the mafic dike swarm area, as shown in Fig. 14. The distribution of estimated nepheline, K-feldspar, biotite, hornblende, clinopyroxene, albite, titanite and magnetite in the deposit are shown in Fig. 15.

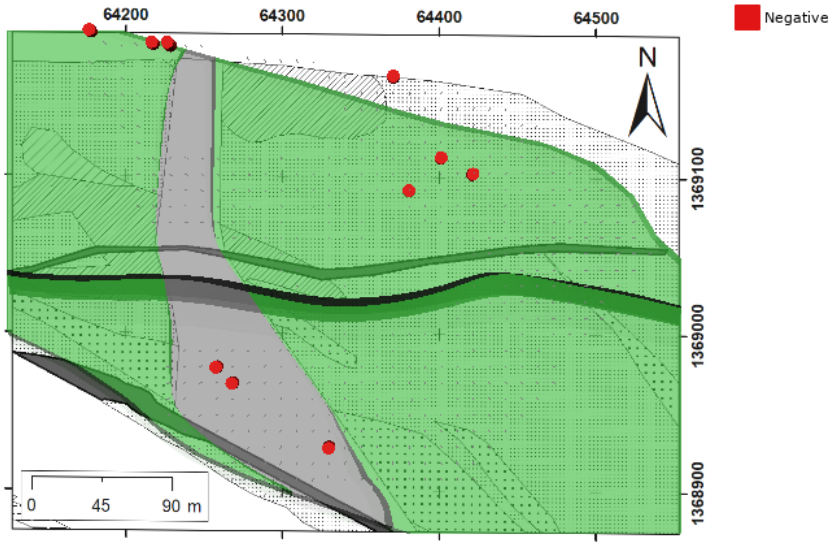

Fig. 14. *Open pit top view of the modal mineralogy estimations from the NN1 model. Negative mineral estimates are displayed in red to illustrate their spatial distribution. Positive mineral estimations are not displayed. The shapes of the fault (black), mafic dike swarm (grey) and mineable nepheline syenite (green) areas vary with deposit depth.*

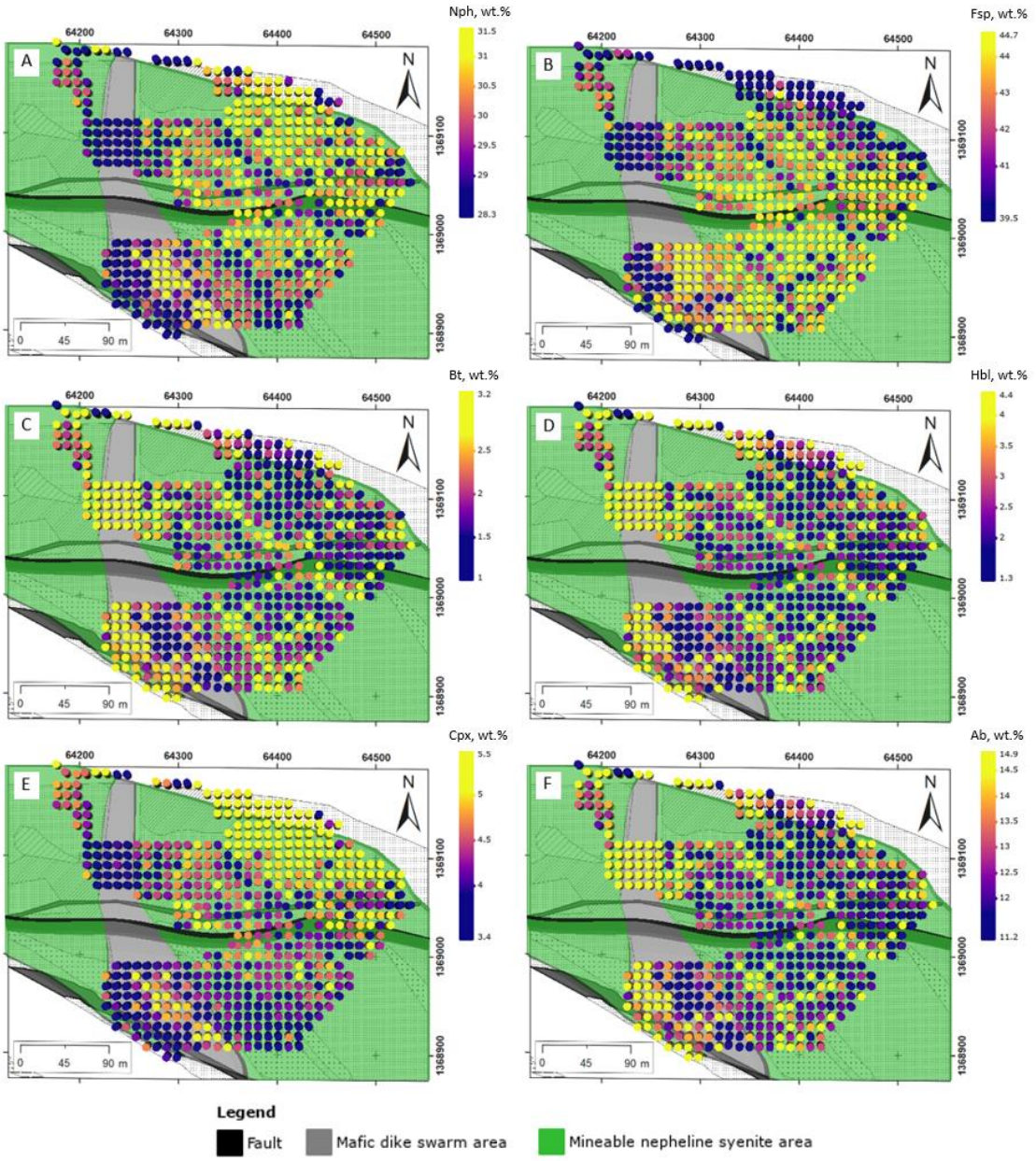

Fig. 15. *Open pit top view of a) nepheline, b) K-feldspar, c) biotite, d) hornblende, e) clinopyroxene, f) albite estimations. The shapes of the fault, mineable nepheline syenite and mafic dike swarm areas vary with deposit depth. The legend in each plot is based on the histogram of every variable split into quartiles: the Q1 (lower quartile or 25th percentile of the data) is in the blue transition, the Q2 (median or 50th percentile) is in the purple transition, and the Q3 (upper quartile or 75th percentile) is in the yellow transition. The extreme values in the legend are a representation of the tails from each element histogram and correspond to the Q1 and Q3 values expanded by 0.5 wt.%.*

4.6. Concentrate yield estimations and comparison

The comparisons between measured and estimated concentrate yield from the three neural network models are shown in Fig. 16. The spatial distributions of the measured and estimated concentrate yield by model in the deposit are shown in Fig. 17.

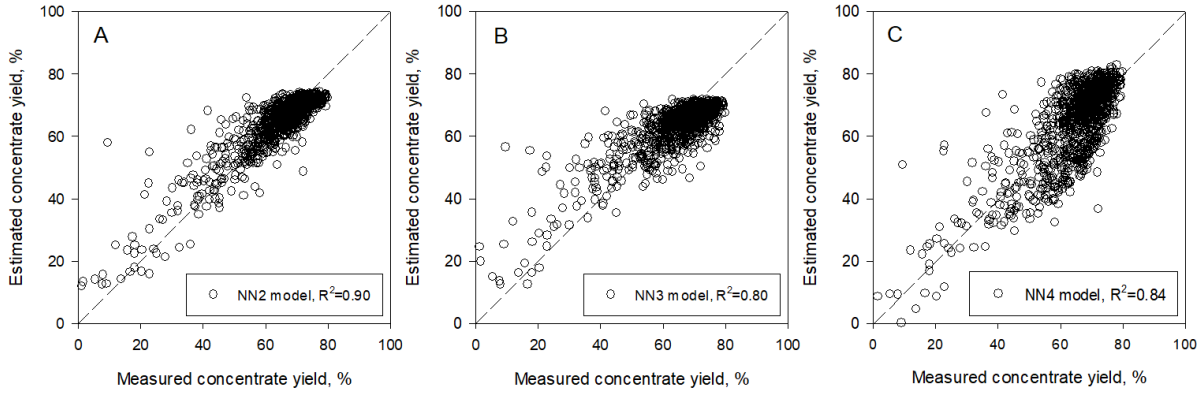

Fig. 16. *Concentrate yield prediction performance applied to the company bulk chemical database by model a) NN2 model with bulk chemistry and location class as input, b) NN3 model with estimated mineralogy and location class as input, and c) NN4 model with bulk chemistry, estimated mineralogy and location class as input.*

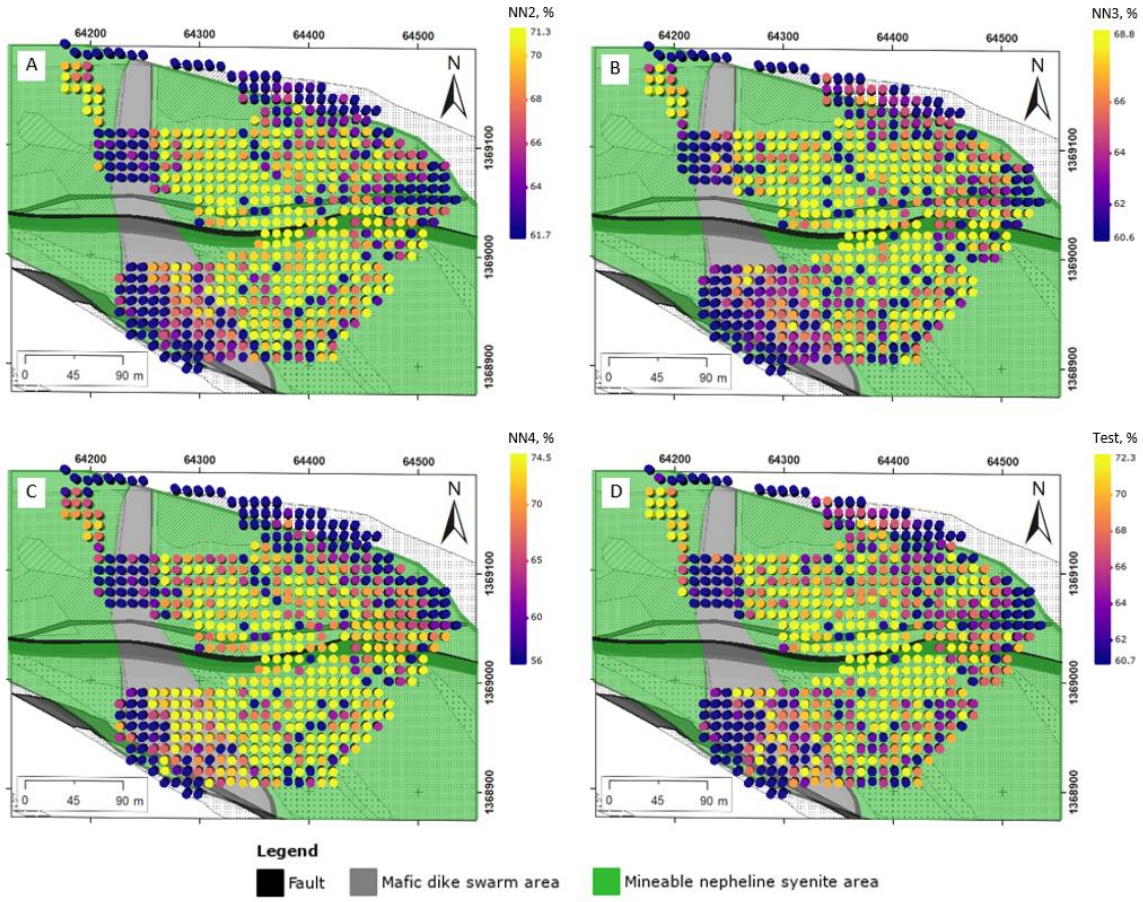

Fig. 17. *Open pit top view of the concentrate yield distribution by a) NN2 model prediction, b) NN3 model prediction, c) NN4 model prediction, and d) measured concentrate yield. The shapes of the fault, deposit production and mafic dike swarm areas vary with deposit depth. The legend in each plot is based on the histogram of every variable split into quartiles: the Q1 (lower quartile or 25th percentile of the data) is in the blue transition, the Q2 (median or 50th percentile) is in the purple transition, and the Q3 (upper quartile or 75th percentile) is in the yellow transition. The extreme values in the legend are a representation of the tails from each element histogram and correspond to the Q1 and Q3 values expanded by 0.5 wt.%.*

The concentrate yield prediction model NN2 had no negative estimations and a *mse* of 0.16 on data not used to train the model This is considered as the best concentrate yield prediction model. The indicator group system from Table 4 was applied to the NN2 model concentrate yield predictions for further evaluation as shown in Fig. 18. The +4sd indicator

group has the highest differences between measured and estimated data. The points belonging to this group are mainly located in the periphery of the deposit, in the mafic dike location class and the proximities of fault crossing the deposit, whereas the other groups are randomly distributed in the whole deposit as shown in Fig. 19.

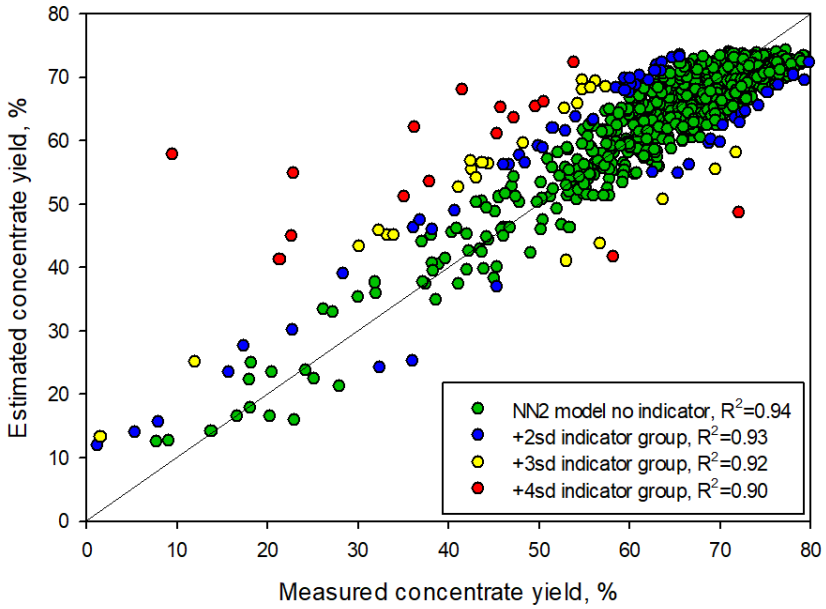

Fig. 18. *Concentrate yield correlation between measured and NN2 model estimated values, colour code according to indicator group and its influence in the overall correlation coefficient.*

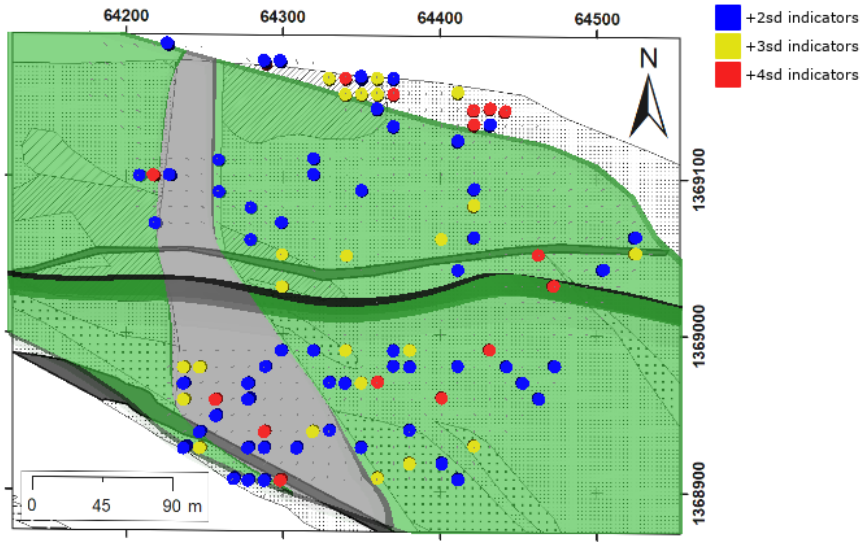

Fig. 19. *Open pit top view of the +2ds, +3ds and +4ds indicator groups from the NN2 model concentrate yield estimations. The shapes of the fault (black), mafic dike swarm (grey) and mineable nepheline syenite (green) areas vary with deposit depth.*

The summary statistics of the three indicator groups are displayed in Table 11. There are differences between each indicator group such as Al_2O_3, Fe_2O_3, and concentrate yield obtained with them, though other variables like estimated mineralogy do not show any particular differences.

Table 11 *Summary statistics of the indicator groups from the NN2 concentrate yield model. Nph = nepheline, Fsp = K-feldspar, Ab = albite, Cpx = clinopyroxene, Hbl = hornblende, Bt = biotite, Ttn = titanite, Mgt = magnetite, Po = pyrrhotite, Cal = calcite, Ntr = natrolite, based on Whitney and Evans (2010) nomenclature.*

Indicator group	**+2sd**			**+3sd**			**+4sd**		
Points	59			25			16		
Variable, wt.%	Mean	Std. dev.	Skewness	Mean	Std. dev.	Skewness	Mean	Std. dev.	Skewness
Al_2O_3	21.32	2.41	-1.61	20.84	2.42	-0.94	21.52	1.67	-0.51
BaO	0.31	0.09	0.02	0.35	0.09	-0.33	0.38	0.07	0.31
CaO	3.69	2.58	3.63	3.98	1.48	1.41	3.38	0.85	0.44
K_2O	7.56	1.57	-1.85	7.36	1.16	-3.08	7.61	0.64	0.26
MgO	1.40	1.90	2.57	1.80	2.17	2.69	1.12	0.76	0.53
Na_2O	6.93	1.09	-1.16	6.34	1.25	-1.21	6.92	1.19	-1.08
SiO_2	52.50	2.98	-2.68	52.13	2.02	-0.80	52.84	1.35	0.12
Fe_2O_3	3.83	2.13	1.48	4.26	1.90	1.42	3.86	1.33	0.73
TiO_2	0.85	0.54	2.94	0.99	0.43	1.45	0.89	0.31	0.53
Concentrate yield %	54.65	19.24	-1.11	45.96	16.13	-1.02	40.59	15.87	-0.15
Nph	29.14	2.65	-0.10	28.27	2.24	-0.26	28.88	2.90	-0.19
Fsp	40.85	4.05	-0.52	39.61	3.83	0.21	39.92	4.18	-0.11
Ab	14.32	3.58	0.53	15.52	3.12	0.23	14.90	3.74	0.56
Cpx	4.48	0.97	1.76	4.30	0.91	0.63	4.58	1.23	1.05
Hbl	3.87	2.77	0.56	4.79	2.43	0.20	4.34	2.88	0.56
Bt	2.61	1.53	0.35	3.12	1.30	0.30	2.81	1.63	0.43
Ttn	0.39	0.15	1.85	0.37	0.15	0.60	0.41	0.19	1.02
Mgt	0.99	0.10	-0.40	0.96	0.10	0.38	0.97	0.11	0.26
Po	0.29	0.09	0.79	0.26	0.08	0.18	0.29	0.10	0.56
Cal	2.00	0.17	1.84	1.97	0.17	0.62	2.02	0.22	1.04
Ntr	1.15	0.53	-0.51	0.99	0.51	0.24	1.03	0.55	-0.07

5. Discussion

5.1. Uncertainty in the data

The modal mineralogy is assumed to be known from the XRD analyses, though there is no benchmark to compare against. Hence the assumed *known* mineralogy relied upon one technique. However, there are uncertainties associated to XRD analyses like the high detection limit where minerals quantities lower than 0.5-1.0 wt.% such as apatite and corundum have low precision, if detected at all, in bulk rock analyses. The methodology in the present study, used a predefined mineral list during Rietveld refinement that normalizes the minerals present in the samples and forces the analyses to fit 100 wt.%, thus neglecting the presence of minerals not considered in the mineral list.

Despite the limitations mentioned, it is believed that the effect of the uncertainties will not have a large impact on the modal mineralogy estimation results nor the general trends observed from the estimated modal mineralogy in the deposit. Also, the use of a standardized procedure, as the one used in the present study, is desired towards the automation of the XRD refinement.

Other aspects to consider is associated to the EPMA analyses used during Rietveld refinements. These analyses are also prone to subtle uncertainties and limitations that should be mentioned such as the inability to detect the lightest elements i.e. H and thus hydrous minerals, nor to differentiate element valence states e.g. Fe^{2+} or Fe^{3+} as described by Goldstein et al. (2003). Additionally, this technique is prone to underestimate light elements due to mobilisation, which can be induced by the electron beam during analyses as proven by Szalóki et al. (2001). Despite this, the use of EPMA analyses on the deposit material is preferred over standard mineral chemistry values from software packages because the EPMA mineral chemistry data represents better the minerals present in the Nabbaren deposit.

Expanding the list of features to characterize the material is desirable and it could be relatively fast and cost efficient. Some examples are the use of loss-on-ignition (LOI), a major element list expansion i.e. by adding carbon, sulphur, etc., or minor element measurements. The LOI indicates the presence of volatile components in the material such as carbonates, hydrous minerals or sulphides; minor elements can provide proof of the presence of a certain mineral (i.e. similarly to the assumption of phosphorus to apatite). These extra measurements can improve the company database and could be included into the internal procedures.

Other automated techniques, such as SEM automated mineralogy software, with lower detection limits and the capability to quantify other sample features (e.g. liberation degree, mineral associations, etc.) provide extra data to improve knowledge that can have a positive impact in the predictions. Nonetheless, this type of technique could struggle with perthitic textures (Hestnes and Sørensen, 2012; Smith and Brown, 1988), such as those from the Nabbaren deposit.

5.2. Variability introduced from the acquisition method

Previous geological investigations (Geis, 1979; Heier, 1961; Li, 2013; Oosterom, 1963; Robins, 1980; Robins and Often, 1996; Robins and Tysseland, 1979) indicate that amphibole-rich rocks like a syenitic fenite dominate southwest of the Nabbaren deposit, while carbonatites and nepheline-rich gneisses are dominant in the northeast. Observations from the company also explain part of the measured data variation with the observation of an area swarmed with diabase dikes crossing from north to south in the deposit and the faults, as shown in Fig. 4.

The measured data of concentrate yield (Fig. 16 subplot D) and the bulk chemistry (Fig. 5) in the deposit could be influenced by the lithologies surrounding the deposit as illustrated in Fig. 2, since the sampling was not based on a map with a clear distinction between the nepheline syenite and surrounding lithologies, or of internal variations in the nepheline syenite.

An aspect introducing variability in the data that could affect parameter correlations and the outcome of the prediction models could be the reverse-circulation drilling sample acquisition method. These samples do not represent a *pure* sample; on the contrary, the resulting samples might contain a mixture of lithologies likely from hanging-wall or footwall mixing in the periphery or mixed lithologies in the mafic dike swarm and fault. The two location classes defined in Fig. 6, the mafic dike swarm and the nepheline syenite areas, are both a mixture of mafic dikes and nepheline syenite. A lithological mixture can introduce unwanted variation to the data, and therefore make any prediction and interpretation harder to achieve. Still it is expected that underlying chemical and mineralogical patterns are detectable also when using this sample type. The reverse-circulation sample is the current sample type used by the company and thus comprising the whole database available for the deposit. It is reasonable for the company to use the samples available for a more in-depth characterization despite the problems associated with them.

The presence of moderate amounts of mafic dikes should not impact the concentrate from magnetic separation according to Geis (1979), though the same cannot be stated about the impact on concentrate yield which evidently varies between the measured points located in the mafic dike swarm area, the nepheline syenite deposit and the rim of the deposit and shown in Fig. 16 (subplot D).

5.3. Data correlation and model confirmation

The correlations between measured data showed a relation between modal mineralogy, bulk chemistry and concentrate yield. Some of the highest correlation values from Table 6, Table 7 and Table 8 are observed between concentrate yield and bulk chemistry, such as K_2O (positive), TiO_2, CaO and Fe_2O_3 (negative), and that between mineralogy and bulk chemistry like nepheline with MgO (negative) and Na_2O (positive), K-feldspar with K_2O, Al_2O_3 and SiO_2 (positive), hornblende with K_2O (negative), MgO, TiO_2 and Fe_2O_3 (positive), biotite with MgO (positive) and titanite with BaO (positive).

The Pearson's coefficient from Table 6 might indicate a linear type of relation between sets. However, this relation is not entirely true, as this coefficient is highly influenced by the presence of extreme values in the data. In most variables, Kendall and Spearman coefficients showed similar or higher coefficients than Pearson's. This indicates that the relation between modal mineralogy, bulk chemistry and concentrate yield might be better suited for a non-linear prediction models such as a neural network.

Unfortunately, often the relation between datasets cannot be explained only by modal mineralogy and bulk chemistry. There are multiple influencing factors such as mineral textures (Lund et al., 2015; Tøgersen et al., 2018) mineral associations and the degree of liberation (Hunt et al., 2011; Johnson et al., 2007; Minz et al., 2013). These properties were not measured in this study, though they should be considered in future geometallurgical investigations due to the impact they might have in characterizing the material and its processability (Lamberg, 2011; Williams, 2013).

5.4. The prediction models

The performance of the prediction models is reflected in the obtained *mse*, which are associated to the selected model structure. Less complex structures such as a one-hidden layer with an *optimal* number of neurons tend to predict any function (Demuth et al., 2014; Maier et al., 2000), in this case bulk chemistry to concentrate yield and modal mineralogy.

The predictions for the different minerals, had different performances (Fig. 13). Most of the minerals predicted by the NN1 model seem to fall on the 1-to-1 line, except for clinopyroxene, which showed a rather biased prediction of the measured data. Many factors can influence the model bias such as a lack of variability shown in the data as displayed in Table 5, with clinopyroxene skewness close to zero, a biased training set from the allocation of samples after clustering, an optimal model found after training failed to find a better solution without decreasing the overall performance of the model thus compromising clinopyroxene predictions, a low number of available samples (diabase dike swarm area had 9 samples), just to mention some reasons.

Despite the possible biases from the NN1 model, when applying the model to the bulk chemical database of the deposit, only 9 points predicted negative mineralogy i.e. at least one negative mineral for the sample. Fig. 14 shows the location of these 9 points, which could be linked to the diabase dike swarm area or to the nepheline syenite rim. The negative predictions in these points can be expected due to an extrapolation of the model to samples different from the training set. This is a consequence of the area covered by the sample used for modal mineralogy studies (see Fig. 4) and the data variability (see Table 2 and Table 5).

The prediction of modal mineralogy could suffer from lack of sample representativeness in parts of the location areas and surrounding rock, due to the spatial constraints of the data set, and limited number of samples in some areas.

One way to deal with the lack of representativeness is by measuring extra samples for a re-evaluation and confirmation of the models, which can be done by the company in further investigations. For example, by studying samples at different locations in the open pit or by re-sampling using a more compact sampling grid over certain areas. Alternatively, the use of artificial samples based on original dataset statistics could be an option to increase the number of points to create the prediction model as done by Lachtermacher and Fuller (1994). However, an accurate modal mineralogy model is not so critical for the mafic dike area of the present study. The most important aspect with distinguishing this location class was to remove the impact of the mafic dikes on the variations in the main nepheline syenite deposit.

The concentrate yield prediction models did not have the same lack of representativeness problem as the modal mineralogy prediction models. The database consisting of 1055 measured points in the open pit was large enough to cluster and representatively allocate concentrate yield points into the different subsets prior to modelling, as illustrated in Fig. 12. The prediction from the NN2, NN3 and NN4 models (Fig. 16) showed that the best performing model is the NN2 (Table 10). The distribution of the predicted and measured concentrate yield points in the open pit (Fig. 17) show the differences between models. The model input data had certainly impacted the model prediction performance like in the NN3 model, the worst performing model, which used as input data of the estimated modal mineralogy from the NN1 model. The NN2 model only used measured bulk chemical data to estimate concentrate yield and this was the best performing model. Despite the performance shown by the concentrate yield prediction models, the results might indicate a risk of overestimation that has to be considered in a future application of this type of model.

The processing behaviour, such as the magnetic separation concentrate yield, may be explained using bulk chemical data, although mineralogical data is generally regarded as more suitable for this type of task (Evans et al., 2011; Lotter et al., 2011; Lund and Lamberg, 2014; Petruk, 2000). Mineralogical information should provide a more meaningful input to the processing prediction; however, the input data should be measured instead of estimated. Nonetheless, the use of bulk chemical data is a good starting point for a prediction model. Further mineralogical characterization might help understanding the underlying processes involving and influencing processing behaviour. This type of information should be considered by the company in the future development of the mine, processing plant and any prediction model attempt.

When studying further the prediction accuracy of the NN2 model it is expected for the data regarded as outliers to be somehow influenced by their location. In general, the further the estimation is from the centre of the deposit, the further the estimation is from the one-to-one line in Fig. 18. The impact of the outliers or indicator groups in the overall correlation between measured and estimated concentrate yield is particularly interesting because these groups influence the correlation between measured and predicted points. Removing these more extreme values increases the correlation from 0.90 to 0.94. Most of the outliers belong to the mafic dike swarm area, the rim of the nepheline syenite and the fault, as shown in Fig. 19. These locations are most likely heavily influenced by lithologies not well represented in the dataset due to the location of the measured sample points and a possible lithological mixing, as previously mentioned.

The model structure like a one-hidden layer neural network is another aspect to consider. This type of structures can predict basically any output, though the predictions can be improved by using more complex structures (Maier et al., 2000; Mhaskar and Poggio, 2016). The compromise between the complexity of the model and the predicting performance is a

subject under constant development in the field of deep learning (Bishop, 2016; Goodfellow et al., 2016).

The neural network methodology implemented for the prediction of concentrate yield and modal mineralogy show potential for improvement and implementation in operations for fast-track data. For example, the use of predicted modal mineralogy can be of great use for the interpretation of any underlying geological information that otherwise would be invisible with bulk chemical data. This type of information can help differencing and interpreting geological features, which has potential in mine planning or resource assessment e.g. defining a meaningful classification system for the deposit based on the database. The predicted concentrate yield has even more potential because the model can be implemented for fast-tracking results that otherwise would take an extensive amount of time in the laboratory. This type of information has also the potential to be used in blending strategies, as a plant forecasting performance indicator, and as a resource assessment indicator for different production levels e.g. to implement geometrical models of the deposit.

5.5. Trends revealed from the prediction models NN1 and NN2

To be of use in the quality control and raw material blending of an industrial mineral deposit, the modal mineralogy and concentrate yield prediction models do not need to accurately predict all samples, but they need to be close on average on a blast scale and reflect the real trends in the deposit. In addition, they might reveal other aspects that could be examined by further studies to improve the knowledge of the quality parameters of the deposit.

The modelled variation in modal mineralogy of the nepheline deposit with surrounding areas is a function of the mineralogical variation of the nepheline syenite (Heier, 1961; Mjelde, 1983; Robins and Gardner, 1975), the mineral chemistry variation of the deposit (Heier, 1964; Mena Silva et al., 2018; Mjelde, 1983), the surrounding lithologies in samples located in the proximities of the deposit rim (Appleyard, 1974; Geis, 1979; Heier, 1961; Robins, 1980; Skogen, 1980), and the number of samples analysed in the different areas for the present study.

In spite of the potential representativity problems associated with the modal mineralogy prediction model due to the sample locations and low number of modal mineralogy analyses, the modal mineralogy predictions in the deposit (Fig. 15) showed interesting results in accordance with the observed bulk chemical variation and the lithologies surrounding the deposit. The variation in the deposit showed a differentiation between the northeastern and southwestern parts of the deposit. The northeastern part of the deposit had higher concentrations of nepheline, clinopyroxene and titanite, while the southwestern part of the deposit had higher concentrations of K-feldspar, albite, biotite, and magnetite. The differentiation between the distributions of biotite and clinopyroxene is expected from previous deposit descriptions (Geis, 1979; Heier, 1961). The association between clinopyroxene and titanite has been highlighted in the first descriptions of the deposit by Heier (1961) that reported a positive correlation between these minerals.

The representativity problem of the modal mineralogy prediction model is especially important in areas with mixed lithologies, as at the rims of the deposit and in the mafic dike swarm area.

The nepheline syenite area is in contact with different lithologies such as carbonatite and nepheline-rich gneiss in the northeast, and syenitic fenite in the southwest. Considering the sample location, the modal mineralogy of these rim areas might be influenced by the location and number of samples analysed. For instance in Fig. 4, due to the mineralogical

characteristics of carbonatite and nepheline-rich gneiss (Geis, 1979; Heier, 1961; Oosterom, 1963; Robins and Gardner, 1975; Skogen, 1980), they can be considered as the main lithologies mixing in the northern part of the nepheline syenite deposit, influencing the estimated calcite and nepheline concentrations. In the southwestern part of the deposit and the diabase dike areas with fewer modal mineralogy analyses, estimated biotite, plagioclase and K-feldspar might be influenced by the contacting syenitic fenite located in the southern part of the deposit (Geis, 1979; Heier, 1961; Ramsay and Sturt, 1970).

The measured modal mineralogy showed a difference between the nepheline syenite and mafic dike swarm areas, as displayed in Table 5. The same is reflected in the predicted modal mineralogy model. The nepheline syenite area had higher estimated concentrations of nepheline, clinopyroxene and calcite, whereas the diabase dike swarm area had higher concentrations of K-feldspar, albite, and biotite.

In addition to this, the modelled mineralogy of the mafic dike area showed a spatial differentiation. In the south the mafic dike area splits between a western and an eastern part. The samples belonging to the western part show similar values to the ones located in the north of the mafic dike swarm, while samples belonging to the eastern part show values similar to what is expected in the nepheline syenite area. This shows the true nature of the mafic dike area, as an area with varying densities of mafic dykes cutting the nepheline syenite, unlike the more schematic representation of the mafic dike swarm area shown in Fig. 6. A more detailed mapping of this area could improve the potential for selective mining of this area.

The concentrate yield distribution in the deposit portrays a general variation of the modal mineralogy and bulk chemical data in the deposit and the different areas within it. The mafic dike swarm area has higher concentrations of CaO, MgO, Fe_2O_3 and a lower concentrate yield, whereas the nepheline syenite area has higher concentrations of Al_2O_3, K_2O, Na_2O and a higher concentrate yield (Table 2).

The estimated concentrate yield is in general in very good agreement with the measured values, suggesting that model NN2 based on bulk chemistry could be used as a proxy for concentrate yield within the investigated area.

The variation in the characteristics from both location classes and the whole dataset indicate that the current company quality map could benefit from a better definition of domain boundaries and understanding of the internal quality variations of the quality domains. This applies to both the nepheline syenite and the mafic dike swarms.

Further work regarding prediction models and deposit characterization should address the modal mineralogy representativity issue, focusing on deposit domaining, and additional mineralogical studies. Additional modal mineralogy measurements covering a larger area in the deposit should be carried out based on a reconciled geological map using the available bulk chemistry database. Evaluating other neural network architectures to increase the precision of the predictions could be a further task to compare performances with simpler structures as the one from the present study, but simple structures seem to give useful results given sufficient data covering the sampled area. The measurement of extra input features (like LOI and mineral texture) towards a more complete material characterization could be considered by the company in future sampling campaigns and operations.

6. Conclusion

The present research aimed at developing and testing prediction models using a neural network framework for mineral processing performance indicators such as concentrate yield

and modal mineralogy based on bulk chemistry. To use bulk chemistry data from existing reverse-circulation samples, was of particular interest because to use this data set would represent the fastest and most cost-efficient way of obtaining information on mineralogy and concentrate for the company.

From the present study it is possible to conclude:

By using a neural network approach a model was obtained which predicts concentrate yield accurately based on bulk chemistry by achieving a correlation coefficient of 0.9 between estimated and measured values.

The relation between concentrate yield, sample bulk chemistry and modal mineralogy is better described in non-linear terms based on the different correlation coefficients tested. From the non-linearity premise, a neural network approach was evaluated to be preferred to predict concentrate yield and modal mineralogy on bulk chemistry.

The estimation of modal mineralogy showed biased predictions, though it is expected that the bias is influenced by the location and number of samples used as input. Nonetheless, the modal mineralogy predictions have potential for improvement if more samples are collected and included during model training.

From the measured data it was possible to recognize a trend from the southwestern to the northeastern part of the deposit, assumed to be related to the variation in mineralogy and mineral chemistry, as well as due to influence from surrounding lithologies and internal fault zone and mafic dikes.

The use of reverse circulation samples has its limitations but is believed to reveal the true trends of the studied deposit.

The distribution of the predicted and measured data in the deposit indicate that the current quality map describing the deposit could benefit from a better definition of domain boundaries and understanding of the internal quality variations of the quality domains, due to inconsistencies between the current map features and the distribution of the measured and predicted data.

The application of a neural network approach showed a successful attempt in the prediction of concentrate yield and modal mineralogy in the Nabbaren nepheline syenite deposit. However, further investigations in terms of deposit internal variation and mineralogical studies are needed for utilising these prediction models, and further improve the modal mineralogy prediction model.

Acknowledgments

Thanks to Torill Sørløkk and Laurentius Tijhuis for helping with the XRF and XRD data acquisition, and to Maarten Felix and Stephen Lippard for English correction and proof reading. Thanks for all the technical support from the Geoscience and Petroleum department at NTNU and Sibelco Nordic AS at the Island of Stjernøy.

This research has been funded by the Research Council of Norway and the industry partners through the InRec project, grant no. 236638. Special thanks to Sibelco Nordic AS for the cooperation.

References

Aasly, K., Ellefmo, S., Geometallurgy applied to industrial minerals operations. Mineralproduksjon, 2014, **5**, 21-34.

Appleyard, E., Syn-orogenic igneous alkaline rocks of eastern Ontario and northern Norway. Lithos, 1974, **7(3)**, 147-169.
Barth, T.F.W., The composition of nepheline. Schweizerische Mineralogische und Petrographische Miteilungen, 1963, **43(1)**, 153-164.
Berry, R., Hunt, J., McKnight, S., 2011. Estimating mineralogy in bulk samples, In *1st International Geometallurgy Conference (GeoMet 2011)*, pp. 153-156.
Berry, R., Hunt, J., Parbhakar-Fox, A., Lottermoser, B., 2015. Prediction of Acid Rock Drainage (ARD) from Calculated Mineralogy, In *10th International Conference on Acid Rock Drainage and IMWA Annual Conference*, pp. 1-10.
Bishop, C., *PATTERN RECOGNITION AND MACHINE LEARNING*. 2016, Springer, New York.
Brough, C., Warrender, R., Bowell, R., Barnes, A., Parbhakar-Fox, A., The process mineralogy of mine wastes. Minerals Engineering, 2013, **52**, 125-135.
Bryan, W.B., Finger, L.t., Chayes, F., Estimating proportions in petrographic mixing equations by least-squares approximation. Science, 1969, **163(3870)**, 926-927.
Cheary, R.W., Coelho, A., A FUNDAMENTAL PARAMETERS APPROACH TO X-RAY LINE-PROFILE FITTING. J. Appl. Crystallogr., 1992, **25**, 109-121.
Cropp, A., Goodall, W., Bradshaw, D., Hunt, J., Berry, R., Communicating and integrating geometallurgical data along the mining value chain. Impc 2014, 2014, 1-12.
Croux, C., Dehon, C., Influence functions of the Spearman and Kendall correlation measures. Statistical Methods and Applications, 2010, **19(4)**, 497-515.
Demuth, H.B., Beale, M.H., De Jesus, O., Hagan, M.T., *Neural network design*. 2014, Martin Hagan.
Dominy, S., O'Connor, L., 2016. Geometallurgy - beyond conception, In *Geometallurgy 2016*. The AusIMM, pp. 3-10.
Dunham, S., Vann, J., 2007. Geometallurgy, geostatistics and project value – does your block model tell you what you need to know?, In *Project Evaluation Conference*, Melbourne, pp. 189-196.
Dunham, S., Vann, J., Coward, S., Beyond Geometallurgy – Gaining Competitive Advantage by Exploiting the Broad View of Geometallurgy. Geomet 2011, 2011**(September)**, 5-7.
Evans, C., Wightman, E., Manlapig, E., Coulter, B., Application of process mineralogy as a tool in sustainable processing. Minerals Engineering, 2011, **24(12)**, 1242-1248.
Gallagher, M., Deacon, P., 2002. Neural networks and the classification of mineralogical samples using X-ray spectra, In *Neural Information Processing, 2002. ICONIP'02. Proceedings of the 9th International Conference on*. IEEE, pp. 2683-2687.
Geis, H.P., Nepheline Syenite on Stjernøy, Northern Norway. Economic Geology, 1979, **74**, 1286-1295.
Goldstein, J.I., Newbury, D.E., Michael, J.R., Ritchie, N.W., Scott, J.H.J., Joy, D.C., *Scanning electron microscopy and X-ray microanalysis*. Third edn. 2003, Springer.
Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y., *Deep learning*. 2016, MIT press Cambridge.
Heier, K.S., Layered gabbro, hornblendite, carbonatite and nepheline syenite on Stjernøy, North Norway. Norsk Geologisk Tidsskrift, 1961, **41**, 155-190.
Heier, K.S., A NOTE ON THE U, Th, AND K CONTENTS IN THE NEPHELINE SYENITE, AND CARBONATITE ON STJERNÖY, NORTH NORWAY. Norsk Geologisk Tidsskrift, 1962, **42(3)**, 287-292.
Heier, K.S., GEOCHEMISTRY OF THE NEPHELINE SYENITE ON STJERNØY, NORTH NORWAY. Norsk Geologisk Tidsskrift, 1964, **44**, 205-215.
Heier, K.S., Some crystallo-chemical relations of nephelines and feldspars on Stjernöy, North Norway. Journal of Petrology, 1966, **7(1)**, 95-113.
Heier, K.S., Taylor, S.R., A Note on the Geochemistry of Alkaline Rocks. Norsk Geologisk Tidsskrift, 1964, **44**.
Hestnes, K., Sørensen, B., Evaluation of quantitative X-ray diffraction for possible use in the quality control of granitic pegmatite in mineral production. Minerals Engineering, 2012, **39**, 239-247.
Hunt, J., Berry, R., Bradshaw, D., Characterising chalcopyrite liberation and flotation potential: Examples from an IOCG deposit. Minerals Engineering, 2011, **24(12)**, 1271-1276.

Hunt, J.A., Berry, R.F., Geological Contributions to Geometallurgy: A Review. Geoscience Canada, 2017, **44(3)**, 103-118.
Ioffe, S., Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, In *ICMI*.
Johnson, L., Chu, C., Hussey, G., Quantitative clay mineral analysis using simultaneous linear equations. Clays Clay Miner, 1985, **33(2)**, 107-117.
Johnson, R.C., Scott, G.W., Lukey, H.M., Implications of Mineralogy , Grain Size and Texture on Liberation and Pellet Quality of Great Lakes Iron Ore. Iron Ore Conference, 2007**(August)**, 109-111.
Karunanithi, N., Grenney, W.J., Whitley, D., Bovee, K.J.J.o.c.i.c.e., Neural networks for river flow prediction. 1994, **8(2)**, 201-220.
Koch, P.H., Lund, C., Rosenkranz, J., Automated drill core mineralogical characterization method for texture classification and modal mineralogy estimation for geometallurgy. Minerals Engineering, 2019, **136**, 99 - 109.
Koujelev, A., Lui, S.-L., 2011. Artificial Neural Networks for Material Identification, Mineralogy and Analytical Geochemistry Based on Laser-Induced Breakdown Spectroscopy, In *Artificial Neural Networks-Industrial and Control Engineering Applications*. InTech.
Lachtermacher, G., Fuller, J.D., 1994. Backpropagation in hydrological time series forecasting, In *Stochastic and statistical methods in hydrology and environmental engineering*. Springer, pp. 229-242.
Lamberg, P., 2011. PARTICLES – THE BRIDGE BETWEEN GEOLOGY AND METALLURGY, In *Conference in Mineral Engineering*, Luleå, Sweden, pp. 1-16.
Lang, A., Ellefmo, S., Aasly, K., Geometallurgical flowsheet as a tool for designing and communicating geometallurgical programs. Minerals, 2018**(Geometallurgy)**.
Lecun, Y., Bottou, L., Orr, G.B., Müller, K.R., 1998. Efficient BackProp, In *Neural Networks: Tricks of the trade*, eds. Orr, G.B., Müller, K.R. Springer.
Leroy, S., Pirard, E., Mineral recognition of single particles in ore slurry samples by means of multispectral image processing. Minerals Engineering, 2019, **132**, 228 - 237.
Li, X., 2013. Alkaline Magmatism , Water-Rock Interaction and Multiple Metamorphism in the Seiland Igneous Province , Northern Norway, In *Fakultät für Chemie, Pharmazie und Geowissenschaften*. Albert-Ludwigs-Universität Freiburg, p. 200.
Lishchuk, V., Koch, P.H., Lund, C., Lamberg, P., The geometallurgical framework. Malmberget and Mikheevskoye case studies. Mining Science, 2015a, **22(2)**, 57-66.
Lishchuk, V., Lamberg, P., Lund, C., 2015b. Classification of geometallurgical programs based on approach and purpose.
Lishchuk, V., Lund, C., Ghorbani, Y., Evaluation and comparison of different machine-learning methods to integrate sparse process data into a spatial model in geometallurgy. Minerals Engineering, 2019, **134**, 156-165.
Lotter, N., Kormos, L., Oliveira, J., Fragomeni, D., Whiteman, E., Modern Process Mineralogy: Two case studies. Minerals Engineering, 2011, **24(7)**, 638-650.
Lotter, N.O., Baum, W., Reeves, S., Arrué, C., Bradshaw, D.J., The business value of best practice process mineralogy. Minerals Engineering, 2018, **116**, 226-238.
Lund, C., Lamberg, P., Geometallurgy: A tool for better resource efficiency. European Geologist, 2014, **37**, 39-43.
Lund, C., Lamberg, P., Lindberg, T., Development of a geometallurgical framework to quantify mineral textures for process prediction. Minerals Engineering, 2015, **82**, 61-77.
Maier, H.R., Dandy, G.C.J.E.m., software, Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. 2000, **15(1)**, 101-124.
Malvik, T., History and growth of modern process mineralogy. Mineralproduksjon, 2014, **5**, 1-19.
Marquardt, D.W., An algorithm for least-squares estimation of nonlinear parameters. Journal of the society for Industrial Applied Mathematics, 1963, **11(2)**, 431-441.
Masters, T., *Practical neural network recipes in C++*. 1993, Morgan Kaufmann.

Mathworks, 2018a. Global Optimization Toolbox: User's Guide (r2018b): corr.
Mathworks, 2018b. Global Optimization Toolbox: User's Guide (r2018b): trainlm.
Mathworks, 2019. Least-Squares (Model Fitting) Algorithms.
Mekanik, F., Imteaz, M., Gato-Trinidad, S., Elmahdi, A.J.J.o.H., Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. 2013, **503**, 11-21.
Mena Silva, C., Sørensen, B.E., Aasly, K., Ellefmo, S., Geometallurgical Approach to the Element-to-Mineral Conversion for the Nabbaren Nepheline Syenite Deposit. Minerals, 2018, **8(325)**, 20.
Mhaskar, H.N., Poggio, T., Deep vs. shallow networks: An approximation theory perspective. Analysis and Applications, 2016, **14(06)**, 829-848.
Minz, F., Bolin, N.J., Lamberg, P., Wanhainen, C., Detailed characterisation of antimony mineralogy in a geometallurgical context at the Rockliden ore deposit , North-Central Sweden. Minerals Engineering, 2013, **52**, 95-103.
Mjelde, Ø., 1983. Geologi og petrografi av Nabbaren nefelinsyenitt og dens forold til de omliggende syenitter, nefelinsyenittgneis, karbonatitt og fenittiserte gabbroiske bwergarter i sydlige del av Lillebukt alkaline kompleks, Stjernøy., In *Geologisk Institutt Avd. A*. Universitet i Bergen, Bergen.
Mwanga, A., Lamberg, P., Rosenkranz, J., Comminution test method using small drill core samples. Minerals Engineering, 2015, **72**, 129-139.
Oosterom, M.G., The Ultramafites and layered Gabbro Sequences. Leidse Geologische Mededelingen, 1963, **28(1)**, 177-296.
Parian, M., Lamberg, P., Möckel, R., Rosenkranz, J., Analysis of mineral grades for geometallurgy: Combined element-to-mineral conversion and quantitative X-ray diffraction. Minerals Engineering, 2015, **82**, 25-35.
Pérez-Barnuevo, L., Lévesque, S., Bazin, C., Drill core texture as geometallurgical indicator for the Mont-Wright iron ore deposit (Quebec, Canada). Minerals Engineering, 2018, **122**, 130 - 141.
Petruk, W., *Applied Mineralogy in the Mining Industry*. First Ed. edn. 2000, Elsevier Science B.V.
Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., *NUMERICAL RECIPES - The Art of Scientific Computing*. Third edn. 2007, Cambridge University Press, Cambridge.
Rajabinasab, B., Asghari, O., Geometallurgical Domaining by Cluster Analysis: Iron Ore Deposit Case Study. Natural Resources Research, 2018, 1-20.
Ramsay, D.M., Sturt, B.A., The emplacement and metamorphism of a syn-orogenic dike swarm from Stjernoy, northwest Norway. American Journal of Science, 1970, **268**, 264-286.
Rigol-Sanchez, J., Chica-Olmo, M., Abarca-Hernandez, F., Artificial neural networks as a tool for mineral potential mapping with GIS. International Journal of Remote Sensing, 2003, **24(5)**, 1151-1156.
Robins, B., The evolution of the Lillebukt alkaline complex, Stjernøy, Norway. Lithos, 1980, **13(2)**, 219-220.
Robins, B., Gardner, P., The magmatic evolution of the Seiland Province, and Caledonian plate boundaries in northern Norway. Earth Planetary Science Letters, 1975, **26(2)**, 167-178.
Robins, B., Often, M., 1996. The Seiland Igneous Province, North Norway. Field Trip Guidebook, IGCP project 336, pp. 34-34.
Robins, B., Tysseland, M., Fenitization of Some Mafic Igneous Rocks in the Seiland Province, Northern Norway. Norsk Geologisk Tidsskrift, 1979, **59(1)**, 1-23.
Rosa, D., Rajavuori, L., Korteniemi, J., Wortley, M., 2014. Geometallurgical Modelling and Ore Tracking at Kittila Mine, pp. 24-26.
Rozel, A., Clénet, H., Douté, S., Quantin, C., 2014. Mineralogical characterization using neural networks: Composition of mafic minerals in martian meteorites from their spectra, In *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2014 6th Workshop on*. IEEE, pp. 1-4.
Schapiro, N., Mallio, W.L., Park, W.C., 1981. PROCESS MINERALOGY IN ORE DEPOSIT DEVELOPMENT, In *PROCESS MINERALOGY Extractive Metallurgy, Mineral Exploration, Energy Resources*, eds.

Hausen, D.M., Park, W.C. The Metallurgical Society and American Institute of Mining, Metallurgical, and Petroleum Engineers, New York, pp. 25-30.
Skogen, J.H., The structural evolution of the Lillebukt carbonatite, Stjernøy, Norway. Lithos, 1980, **13**, 221-221.
Smith, J.V., Brown, W.L., 1988. Intimate Feldspar Intergrowths In *Feldspar Minerals Volume 1: Crystal Structures, Physical, Chemical and Microtextural Properties* Second Revised and Extended Edition ed. Springer-Verlag, pp. 555-625.
Suazo, C., Kracht, W., Alruiz, O., Geometallurgical modelling of the Collahuasi flotation circuit. Minerals Engineering, 2010, **23(2)**, 137-142.
Szalóki, I., Osán, J., Worobiec, A., De Hoog, J., Van Grieken, R., Optimization of experimental conditions of thin-window EPMA for light-element analysis of individual environmental particles. X-Ray Spectrometry: An International Journal, 2001, **30(3)**, 143-155.
Sørensen, B.E., 2016. Mineralogical investigation of nepheline syenite from Stjernøy. Norwegian University of Science and Technology, Trondheim, p. 18.
Tsuji, T., Yamaguchi, H., Ishii, T., Matsuoka, T., Mineral classification from quantitative X-ray maps using neural network: Application to volcanic rocks. Island Arc, 2010, **19(1)**, 105-119.
Tøgersen, M.K., Kleiv, R.A., Ellefmo, S., Aasly, K.J.M.E., Mineralogy and texture of the Storforshei iron formation, and their effect on grindability. 2018, **125**, 176-189.
Walters, S., Initiatives to Support Geometallurgical Mapping and Modelling. 1st AusIMM International Geometallurgy Conference (GeoMet 2011), 2011**(September)**, 273-278.
Ward, J.H., Hierarchical grouping to optimize an objective function. Journal of the American statistical association, 1963, **58(301)**, 236-244.
Whiten, B., Calculation of mineral composition from chemical assays. Mineral Processing and extractive metallurgy review, 2007, **29(2)**, 83-97.
Whitney, D.L., Evans, B.W.J.A.m., Abbreviations for names of rock-forming minerals. 2010, **95(1)**, 185-187.
Williams, S., A Historical Perspective of the Application and Success of Geometallurgical Methodologies. The second AUSIMM international geometallurgy conference / Brisbane, QLD, 30 September - 2 October 2013 37, 2013**(October)**, 37-47.
Yvon, J., Baudracco, J., Cases, J., Weiss, J., Eléments de minéralogie quantitative en micro-analyse des argiles. Matériaux Argileux, Structures, Propriétés et Applications, 1990, 473-489.