Natural Resources Canada | Ressources naturelles Canada

Digital Accelerator

Natural Resource Canada's Digital Accelerator

Vik Pant, PhD
Chief Scientist and Chief Science Advisor, NRCan

Canada

Natural Resources Canada is committed to improving the quality of life of **Canadians** by ensuring the country's abundant natural resources are developed **sustainably**, **competitively** and **inclusively**

The natural resource sector directly and indirectly accounts for…

~17% of Canada's nominal GDP

~1.7 million jobs

exports valued at >$260 billion, comprising ~49% of total merchandise exports

Science & Technology is Central to NRCan's Mandate

NRCan has a unique role as **performer, user, funder and communicator of S&T**

NRCan employs **~ 2000 people in S&T roles** across Canada in laboratories, smaller facilities and research sites to **conduct innovative science and generate new technologies**

Budget of over **$500 million for S&T** to drive innovation and productivity

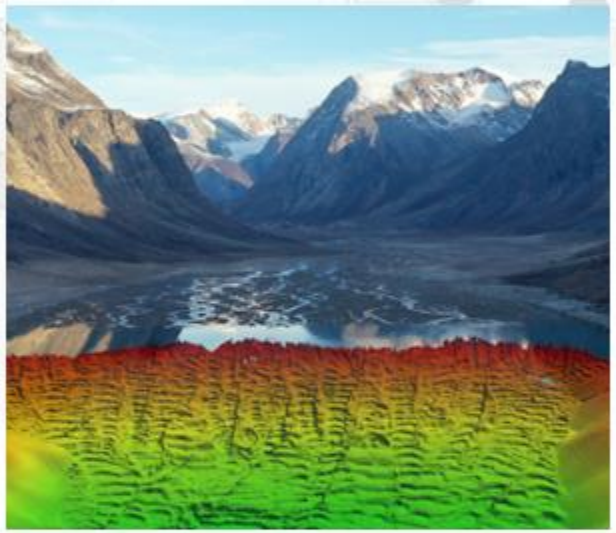
NRCan has a long history of scientific excellence…

Geological Survey of Canada (*Founded in 1842*)

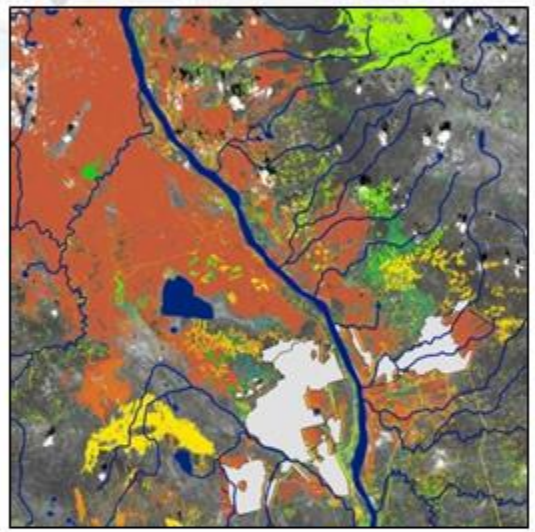
Canadian Forest Service (Founded in 1873 as Dominion Lands Branch)

Canada Centre for Mapping and Earth Observation (Founded in 1970)

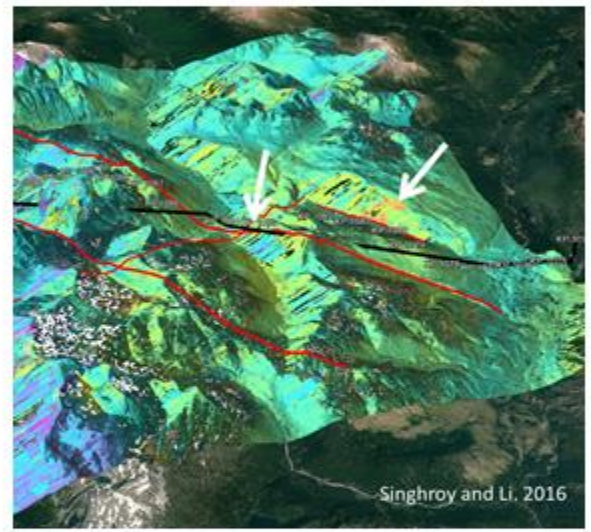
Some Applications of Big Data and Machine Learning



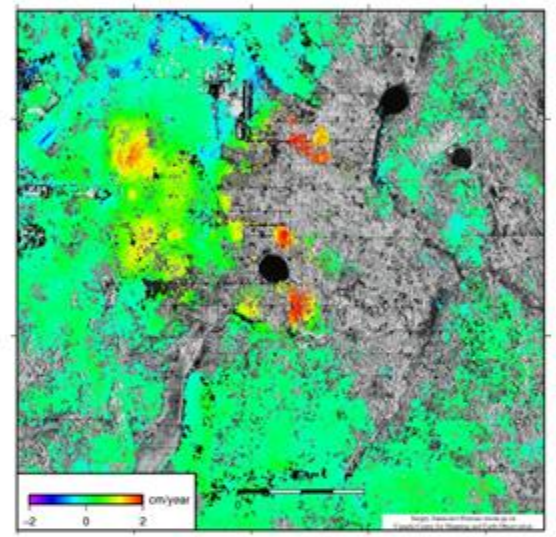
Prediction Spread
of
Groundwater



Estimating Spread
of
Wildfires



Modeling Energy
Yields



Approximating
Infestation
of
Invasive Species

Application of Computer Vision (Machine Learning) in 1990s by CFS

Canadian Journal of Remote Sensing/Journal canadien de télédétection

COMPARISON OF POSSIBLE MULTISPECTRAL CLASSIFICATION SCHEMES FOR TREE CROWNS INDIVIDUALLY DELINEATED ON HIGH SPATIAL RESOLUTION MEIS IMAGES

by FRANÇOIS A. GOUGEON

RÉSUMÉ

Lors de l'analyse par ordinateur d'images aériennes multispectrales de haute définition ayant pour but l'inventaire forestier, il pourrait être plus pratique d'utiliser les cimes d'arbre comme objets de notre analyse plutôt que les peuplements forestiers ou les simples pixels. En partant de cette proposition, il devient important d'être capable de définir spectralement ces cimes le plus succinctement possible. Cet article propose, décrit et compare sept manières différentes de définir spectralement des cimes d'arbre provenant d'images aériennes de haute résolution spatiale dans le but d'identifier leurs espèces forestières.

*Parmis les sept types de signatures spectrales testés, cinq menèrent à des exactitudes de classification du même ordre (72 ± 3 %) en différentiant cinq espèces de conifères. Une amélioration de l'exactitude de classification (jusqu'à 76 %) fut rendu possible pour certainnes signatures en utilisant une procédure d'analyse canonique avant la classification. Les signatures basées sur la «ligne de couleur des arbres», une approche nouvelle, ont donnés des résultats comparable aux signatures plus simples telles que celles basées sur la valeur multispectrale moyenne des cimes d'arbres. Les cimes de pin rouge (**Pinus resinosa** Ait) furent uniformément difficile à séparer spectrallement des autres espèces, alors que celles d'épinette noire (**Picea mariana** (Mill.) B.S.P.) furent uniformément facile.*

SUMMARY

In the computer analysis of high spatial resolution multispectral aerial images for forest inventory purposes, it may be more useful to deal with individual tree crowns as the "objects" of our analysis than with forest stands or individual pixels. Starting from this tenet, it becomes important to be able to spectrally define these tree crowns as succinctly as possible. This paper proposes, describes, and compares seven different ways that tree crowns in high spatial resolution aerial images can be spectrally defined for species classification.

*In testing these seven types of multispectral signatures, it was found that five led to relatively similar classification accuracies (72 ± 3%) in differentiating five coniferous species. Additional classification accuracy improvements (to 76%) were possible with some of the signatures by using Canonical Analysis prior to classification. The "tree colour line"-based signatures, a newly introduced approach, gave comparable results to simpler signatures, such as those based on the mean multispectral value of tree crowns. Red pine (**Pinus resinosa** Ait) crowns were consistently difficult to separate spectrally from other species, whereas black spruce (**Picea mariana** (Mill.) B.S.P.) crowns were consistently easy to distinguish.*

INTRODUCTION

The availability of high-quality georeferenced digital aerial images, such as those produced by the Multidetector Electro-optical Imaging Scanner (MEIS) (McColl *et al.*, 1983), presents an opportunity to re-evaluate the use of digital remote sensing in forestry. Indeed, when subjected to standardized enhancements and used in the context of computer-assisted on-screen image interpretation, such images are likely to change permanently the ways in which forest inventories are done (Leckie, 1990). The easy integration of these images and their georeferenced interpretations

• François Gougeon is with the Petawawa National Forestry Institute, Canadian Forest Service, Natural Resources Canada, Chalk River, Ontario.

Progress in Physical Geography 22,4 (1998) pp. 449–476

Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters

Mike Wulder

Pacific Forestry Centre, Canadian Forest Service, Natural Resources Canada, 506 West Burnside Road, Victoria, BC V8Z 1M5, Canada[1]

Abstract: Forests are the most widely distributed ecosystem on the earth, affecting the lives of most humans daily, either as an economic good or an environmental regulator. As forests are a complex and widely distributed ecosystem, remote sensing provides a valuable means of monitoring them. Remote-sensing instruments allow for the collection of digital data through a range of scales in a synoptic and timely manner. Accordingly, a variety of image-processing techniques have been developed for the estimation of forest inventory and biophysical parameters from remotely sensed images. The use of remotely sensed images allows for the mapping of large areas efficiently and in a digital manner that allows for accuracy assessment and integration with geographic information systems. This article provides a summary of the image-processing methods which may be applied to remotely sensed data for the estimation of forest structural parameters while also acknowledging the various limitations that are presented. Current advancements in remote-sensor technology are increasing the information content of remotely sensed data and resulting in a need for new analysis techniques. These advances in sensor technology are occurring concurrently with changes in forest management practices, requiring detailed measurements intended to enable ecosystem-level management in a sustainable manner.

This review of remote-sensing image analysis techniques, with reference to forest structural parameters, illustrates the dependence between spatial resolution to the level of detail of the parameters which may be extracted from remotely sensed imagery. As a result, the scope of a particular investigation will influence the type of imagery required and the limits to the detail of the parameters that may be estimated. The complexity of parameters that may be extracted can be increased through combinations of image-processing techniques. For example, multitemporal analysis of image radiance values or multispectral image classification maps may be analysed to undertake the assessment of such forest characteristics as area of forest disturbances, forest succession and development, or sustainability of forest management practices. Further, the combination of spectral and spatial information extraction techniques shows promise for increasing the accuracy of estimates of forest inventory and biophysical parameters.

Key words: forest structure; forest inventory; forest biophysical parameters; optical remote sensing; image processing

A method for enhancing tree species proportions from aerial photos

by Steen Magnussen[1]

Data on tree species richness of forests are needed for sustainable management. Photo interpreted data is a mainstay of forest inventories in Canada but is known to simplify the species composition of stands in accordance with the objective of providing stand descriptors. Species proportions estimated from aerial photos will often deviate significantly from ground based sample estimates. More tree species are usually found in the latter. To enhance photo-interpreted species proportions to better match ground observations this study proposes a cross-correlation memory matrix model as the most promising approach to a notoriously difficult problem. Ground and photo data from 246 stands in three sites in coastal BC were used for model estimation and validation. Enhanced photo-based tree species proportions were, in most stands, closer to the ground estimates than the raw estimates. Changes in species proportions due to enhancement appeared reasonable when considering the actual species mix and stand structures. Two popular indices of species richness derived from photo proportions were biased downward. The proposed method effectively reduced the bias in one index. The model is easy to implement in operational inventories.

Key words: photo interpretation, tree species proportions, forest inventories, matrix model

Les données sur la richesse des forêts en terme d'espèces d'arbres sont requises pour un aménagement durable. Les données de photo-interprétation sont à la base des inventaires forestiers au Canada, mais la photo-interprétation est reconnue pour simplifier la composition des espèces formant les peuplements lorsque l'objectif recherché est de reconnaître les éléments descriptifs d'un peuplement. Les proportions des espèces estimées à partir des photos aériennes dévient significativement des estimés établis à partir d'échantillonnages terrestres. On retrouve habituellement plus d'espèces dans ces derniers. Cette étude propose, dans le but de mieux faire concorder les proportions des espèces obtenues par photo-interprétation et au sol, un modèle matriciel à mémoire en corrélation croisée en tant qu'approche la plus prometteuse pour résoudre un problème reconnu comme étant difficile. Les données terrestres et photographiques tirées de 246 peuplements provenant de trois stations de la côte de la C.-B. ont été utilisées pour estimer et valider le modèle. Les proportions améliorées des espèces d'arbres établies à partir des photos ont été, pour la plupart des peuplements, plus près des estimés terrestres que des estimés bruts. Les changements dans les proportions des espèces par suite de l'amélioration apparaissaient raisonnables lorsqu'on considère le mélange actuel d'espèces et la structure des peuplements. Deux indices communs de la richesse des espèces tirés des proportions photographiques étaient biaisés à la baisse. La méthode proposée réduisait effectivement le biais de l'un des indices. Le modèle est facile à implanter lors d'inventaires opérationnelles.

Mots clés: photo-interprétation, proportion d'espèces d'arbres, inventaires forestiers, modèle matriciel

Application of Sequence Learning (Machine Learning) in 1990s by CFS

Modeling Tree-Ring Growth Responses to Climatic Variables Using Artificial Neural Networks

Qi-Bin Zhang, Richard J. Hebda, Qi-Jun Zhang, and René I. Alfaro

ABSTRACT. Modeling the nonlinear and complex relationships between climate and tree-ring growth is of significance in dendroclimatic studies, but difficult to implement using traditional linear regression approaches. To overcome this difficulty, the technique of Artificial Neural Network (ANN) was employed in this study to develop the growth response models using the climate/growth database for Douglas-fir (*Pseudotsuga menziesii* var. *menziesii* [Mirb.] Franco) on southern Vancouver Island, Canada. The results show that the ANN models are able to extract nonlinear growth response patterns from the observed climate/tree-ring datasets, and to generate more accurate predictions than multiple linear regression approaches. The ANN-extracted climate–growth relationships can be displayed by scenario analysis; for example, when all other input variables are held fixed at their means, the limiting effect of April-July precipitation on tree growth decreases with increased precipitation. The main difficulty of applying ANN technique in dendroclimatology is the problem of overlearning (i.e., the ANN learns too many specific climate–growth patterns and loses the ability to generalize between similar climate–growth patterns). This problem can be alleviated by carefully designing the ANN, such as reducing the number of input variables, choosing a variety of training/testing sets, designing partially connected architectures with a small number of neurons in the hidden layer, and using early stopping during training process. The reliability of the derived ANN models is assessed by validation on independent testing datasets. The main advantages of the ANN technique over traditional dendroclimatic approaches are its ability to capture nonlinear climate–growth response, and its nonreliance on preassumed functional relationships for describing the observed datasets. The ANN method introduced in this article is sufficiently general to be applicable to many forest ecological modeling applications. For. Sci. 46(2):229–239.

Additional Key Words: Dendrochronology, dendroclimatology, Douglas-fir, forest growth, response function analysis.

《CHINESE JOURNAL OF APPLIED ECOLOGY》 1999-02

Forest yield prediction with an artificial neural network and multiple regression.

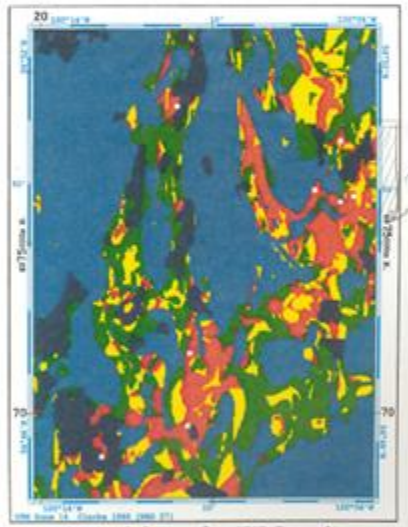
R.Pu, P. Gong (Department of Environmental Science, Policy, and Management, 151 Hilgard Hall, University of California, Berkeley CA94720 3110 USA), R. Yang (Canadian Forest Service, Northwest Region Edmonton, Alberta, Canada T6H 3S5)

Use of traditional statistical techniques is often limited by shortage of observation samples and difference in data measurement scales. Neural network techniques have been extensively explored in many fields for prediction and classification as an alternative to statistical meth ods. In this paper, a feed forward neural network algorithm for predicting hardwood yield is introduced and evaluated. In addition, we report a data transformation method developed for converting qualitative variable data to quantitative data for use in multiple regression when relativ ely few samples are available for building prediction models. The method that converts qualitative data into quantitative data is helpful to imp rove hardwood yield prediction accuracy by multiple linear regression models. In this study, the best prediction results using the neural netw ork technique are obtained.

Application of Machine Learning in 1995 by GSC

Workshop
on
"Concepts and Techniques in Modelling for GIS"

Collection of published and unpublished papers related to modelling in GIS



Potential Map for VMS Deposit
Certainty Factor

Chang-Jo F. Chung
Geological Survey of Canada
601 Booth Street, Ottawa, Canada K1A 0E8
Fax: +613-996-3726; e-mail: chung@gsc.emr.ca

The 7th International Conference on Geomatics
Ottawa, Canada
June 12, 1995

NEURAL NETWORK APPROACH FOR GEOLOGICAL MAPPING: TECHNICAL BACKGROUND AND CASE STUDY

Ping An and Chang-Jo F. Chung

Geological Survey of Canada

601 Booth Street, Ottawa, Canada. K1A 0E8

Fax: +1-613-996-3726; e-mail: chung@gsc.emr.ca

ABSTRACT: The neural network approach for geological mapping can be formulated as approximating a mapping classifier based on a set of training samples and predicting the geological units using the classifier. If the training samples are sufficiently large and the number of hidden nodes of the network are chosen properly, the classification error can be made very small. The number of training samples required for a satisfactory classification depends on the complexity of the problem under investigation. The number of hidden nodes depends on the complexity of the training task which is defined by the training samples. Training coefficients are also critical for successfully approximating the classifier.

The technique is applied to a geological mapping program in Melville Peninsula, Northwest Territories, Canada using 16 data layers including airborne geophysical and satellite-borne LANDSAT TM and SPOT data. Two sets of training samples were selected from the 16 input layers and used for experiments on neural network parameters. One set was based on 510 bedrock pixels (approximately 0.23% of total 221,069 bedrock pixels) and the other was based on 550 outcrop pixels (approximately 1.7% from 32,742 outcrop pixels). The estimated classifiers, when applied to the whole area of 221,069 bedrock pixels and 32,742 outcrop pixels, produced classification predictions with errors less than 3%. The prediction results are also discussed.

Contents lists available at ScienceDirect

Ore Geology Reviews

journal homepage: www.elsevier.com/locate/oregeorev

Predicting rock type and detecting hydrothermal alteration using machine learning and petrophysical properties of the Canadian Malartic ore and host rocks, Pontiac Subprovince, Québec, Canada

Charles L. Bérubé[a,*], Gema R. Olivo[b], Michel Chouteau[a], Stéphane Perrouty[c], Pejman Shamsipour[a], Randolph J. Enkin[d], William A. Morris[e], Leonardo Feltrin[c], Raphaël Thiémonge[a]

[a] Polytechnique Montréal, Département des Génies Civil, Géologique et des Mines, Montréal, QC, Canada
[b] Queen's University, Department of Geological Sciences and Geological Engineering, Kingston, ON, Canada
[c] Laurentian University, Mineral Exploration Research Centre, Harquail School of Earth Sciences, Sudbury, ON, Canada
[d] Geological Survey of
[e] McMaster University,

ARTICLE IN

Keywords:
Petrophysical propert
Supervised machine l
Support vector machi
Mineral exploration
Intrusion-related gold

Journal of Geochemical Exploration 188 (2018) 216–228

Contents lists available at ScienceDirect

Journal of Geochemical Exploration

journal homepage: www.elsevier.com/locate/gexplo

Classification of lithostratigraphic and alteration units from drillhole lithogeochemical data using machine learning: A case study from the Lalor volcanogenic massive sulphide deposit, Snow Lake, Manitoba, Canada

Antoine Caté[a,*], Ernst Schetselaar[b], Patrick Mercier-Langevin[c], Pierre-Simon Ross[a]

[a] Institut national de la recherche scientifique, Centre Eau Terre Environnement, 490 rue de la Couronne, Québec, QC G1K 9A9, Canada
[b] Geological Survey of Canada, 601 Booth Street, Ottawa, ON K1A 0E8, Canada
[c] Geological Survey of Canada, 490 rue de la Couronne, Québec, QC G1K 9A9, Canada

ARTICLE INFO

Keywords:
Lalor
Snow Lake
Mining exploration
Lithogeochemistry
Multivariate classification
Machine learning

ABSTRACT

Classification of rock types using geochemical variables is widely used in geosciences, but most standard classification methods are restricted to the simultaneous use of two or three variables at a time. Machine learning-based methods allow for a multivariate approach to classification problems, potentially increasing classification success rates. Here a series of multivariate machine learning classification algorithms, together with different sets of lithogeochemistry-derived variables, are tested on samples collected at the Lalor Zn-Cu-Au volcanogenic massive sulphide deposit, to discriminate volcanic units and alteration types. Support Vector Machine and Ensemble method algorithms give the best performance on both classification exercises. Untransformed chemical element concentrations with high classification power are the best-performing variables. Classification success rates are equal or better than those obtained using standard classification methods and are satisfactory enough for the use of the resulting predictions for 2D and 3D modelling of geological units.

AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2014WR015452

Key Points:
- CPT/SMR data used to estimate high-resolution 1-D *K* profiles
- Hydro-geophysical data integration using relevance vector machines
- *K* predictions from the learning machine agree with hydraulic tests

Supporting Information:
- Readme
- Data set

Correspondence to:

Predicting hydrofacies and hydraulic conductivity from direct-push data using a data-driven relevance vector machine approach: Motivations, algorithms, and application

Daniel Paradis[1,2], René Lefebvre[2], Erwan Gloaguen[2], and Alfonso Rivera[1]

[1]Geological Survey of Canada, Quebec City, Quebec, Canada, [2]Institut National de la Recherche Scientifique, Centre Eau Terre Environnement, Quebec City, Quebec, Canada

Abstract The spatial heterogeneity of hydraulic conductivity (*K*) exerts a major control on groundwater flow and solute transport. The heterogeneous spatial distribution of *K* can be imaged using indirect geophysical data as long as reliable relations exist to link geophysical data to *K*. This paper presents a nonparametric learning machine approach to predict aquifer *K* from cone penetrometer tests (CPT) coupled with a soil moisture and resistivity probe (SMR) using relevance vector machines (RVMs). The learning machine approach is demonstrated with an application to a heterogeneous unconsolidated littoral aquifer in a

OGS – GSC Collaboration on Regional Groundwater Studies 2014 – 2019: Conclusion

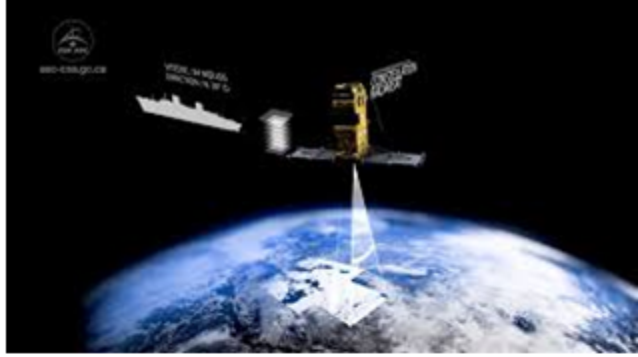
Russell[1], Hazen A.J. and Richard D. Dyer[2]
[1] *Geological Survey of Canada, Natural Resources Canada, Ottawa, ON K1A 0E8*
hazen.russell@canada.ca
[2] *Earth Resources and Geoscience Mapping Section, Ontario Geological Survey, Sudbury, ON P3E 6B5*

The Ontario Geological Survey (OGS) and Geological Survey of Canada (GSC) groundwater collaboration in Southern Ontario will be complete at the end of March 2019. The objective of the collaboration was to maximize technical and human resources to meet objectives of the OGS groundwater initiative and the GSC Groundwater Geoscience Program. To that end the initial collaborative effort involved a GAP analysis in March 2015 which provided guidance for aspects of the subsequent 4 year collaboration. In November 2015 a workshop was held with other provincial ministries, conservation authorities and academia to review a path forward on an improved data framework for sustainable groundwater management. The overall project was orientated along the following principal themes i) Framework for Sustainable Groundwater Use, ii) Supporting Great Lakes Water Accords, iii) Methods Development for Regional Groundwater Studies, iv) 4. Case Studies, v) Science and Technology Exchange.

Recognizing the extensive work completed as part of provincial Source Protection Program the project focused on development of two regional three-dimensional geological models of the bedrock and surficial geology. These models support a fully coupled regional groundwater – surface-water model being developed with Aquanty Inc. The models are supported by OGS, GSC and other provincial, private sector, and academic work and provide a framework for past and future datasets and groundwater understanding. To support this work a number of data collection studies (reflection seismic, downhole geophysics, chemostratigraphic analysis, hydrochemistry) data collation (Municipal well attributes, OGS section data, borehole logs), and data QA and QC activities were completed. To address concerns about groundwater – surface-water in the Great Lakes Water Accord a conceptual framework was developed as a management guide for watershed practitioners. Methods development focused on a

Geo-Deep-Learning

Centre for Mapping and Earth Observation



Challenge: The Canada Centre for Mapping and Earth Observation's Machine Learning model, which classifies geospatial data, requires improvement in terms of processing speed and accuracy

Approach: Improve precision and performance by running it through a Machine Learning testbed

Impacts: Optimized geospatial image interpretation will support decision-making on natural resource development, environmental studies and public safety.

Electric Vehicle Load Prediction

Energy Technology Sector

Challenge: Extensive EV data needs to be modeled to better understand grid loads and support Canada's aim to reduce greenhouse gas emissions

Approach: Take a step-by-step approach to understand and visualize data, and eventually build a predictive Machine Learning model

Impacts: Reduced load on electricity infrastructure by identifying patterns and correlations affecting grid utilization by EVs. Optimized use and management of infrastructure supporting greater use of EVs for diverse transportation needs.

ENERGY STAR® Automation

Low Carbon Energy Sector

Challenge: Misuse and misappropriation of ENERGY STAR® brand by retailers and manufacturers.

Approach: Build a Machine Learning tool to find on the web and flag misuses automatically and aid in enforcement.

Impacts: Strengthening Canadians' ability to make correct energy related decisions.

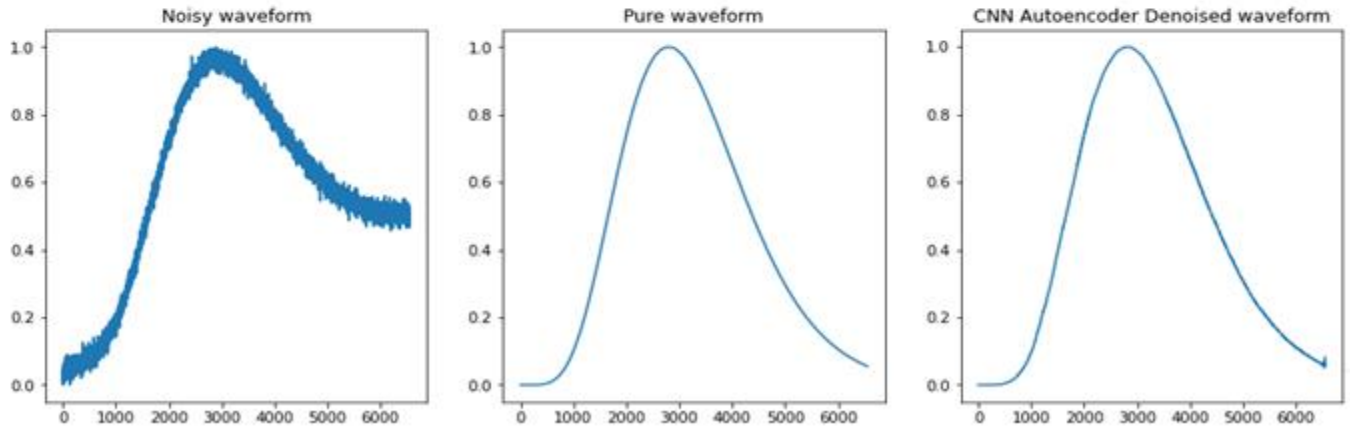
Predictive Particle Modelling

CanmetENERGY

Challenge: To establish baseline operating data to enable construction and operation of the Clean Heat and Power Facility to operate as a system firing natural gas using ilmenite ore as the oxygen carrier.

Approach: Use computational fluid dynamics and machine learning to predict solid circulation rate from sensor inputs.

Impacts: Significant reduction of CO2 emission rates in operations of the Clean Heat and Power Facility, supporting Canada's shift to a low carbon economy.



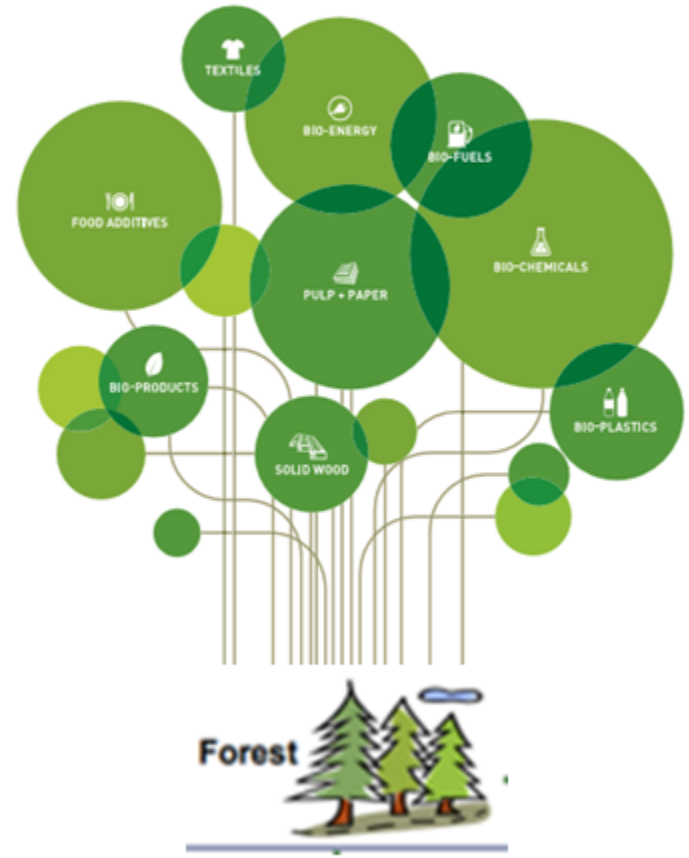
Forest Inventory Monitoring

Canadian Forest Service

Challenge: Accurate species prediction is an enabler in the area of forestry and biomass planning. It can support increasing efficiencies at every step of the forestry value chain including harvesting, processing, scheduling and fostering growth of economy within this sector.

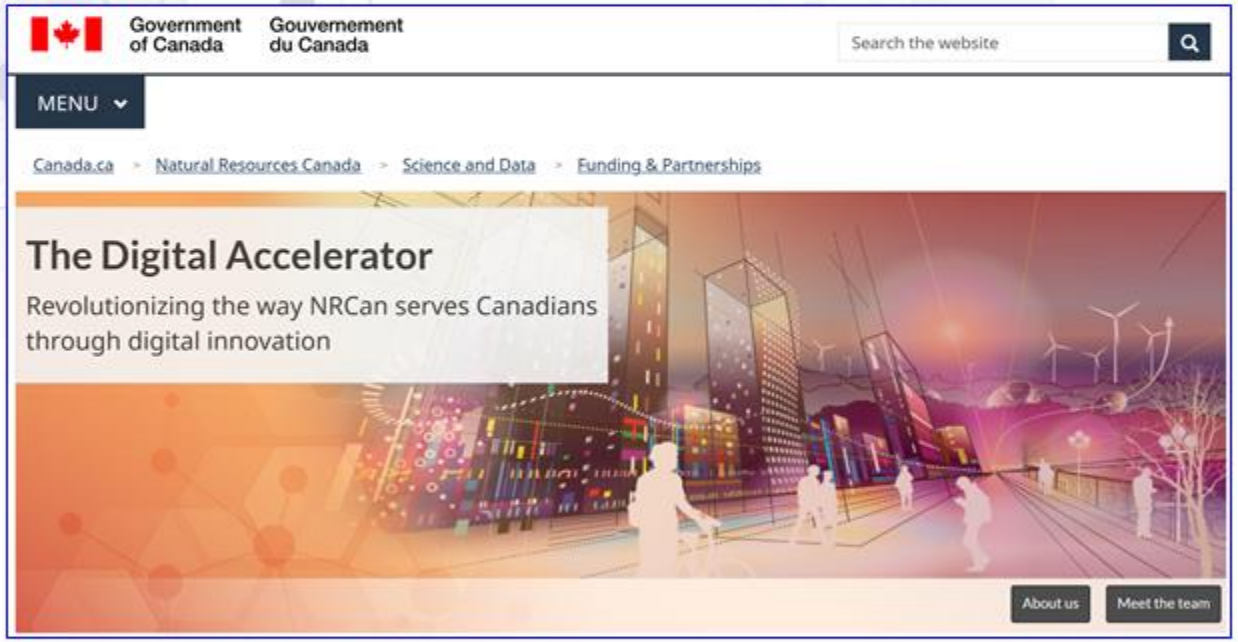
Approach: Use remote sensing data and AI to develop a forest information system that can be used by all stakeholders for forest value chain optimization.

Impacts: Enhanced competitiveness of the sector, stronger support for efficient long and short term planning and strengthened adaptability to deal with climate change impacts.



Thank you

- Contact us at
- https://www.nrcan.gc.ca/digital-accelerator



- nrcan.digital-numerique.rncan@canada.ca