Framework for assessing and communicating uncertainties: Genomic selection for forestry

Resilient Forests (RES-FOR): Climate, Pests & Policy – Genomic Applications

February, 2021

Report compiled by Debra Davidson (University of Alberta) and Gwendolyn Blue (University of Calgary)

“Communicating what is not known is at least as important as communicating what is known.”

Saltelli et al, 2020[i]

Objective and Overview

Communicating and assessing uncertainty is important for decision-making about new technologies, particularly in the face of climate change.

Broadly defined, uncertainty refers to states of incomplete, imperfect, or unknown information. Decisions about forest management often need to be made before conclusive evidence is available. The potential impacts of decisions can be huge. Anticipating issues and discussing potential sources and responses to uncertainty are key to sound decision-making. The decision context matters for the assessment of uncertainty and “one-size-fits-all” approaches do not exist. Decision-makers often need to consider multiple types of uncertainty.

This document provides a framework to guide decision-makers and end-users with assessing and communicating uncertainties associated with genomic selection applications for forestry.

Genomic selection is still ‘in the making’ as a potential technology for forest tree breeding operations. Present and future GS applications will give rise to uncertainties. Government, industry and academic researchers have choices about the extent to which the inherent uncertainties associated with GS are communicated.[ii]

New biotechnologies are on the horizon with potential applications in Canadian forestry and forest management. The Canadian government has made significant investments in the development of **genomic selection** for forestry.

Genomic selection (GS) uses statistical models for making selections (e.g. for seed orchards and breeding) based on associations between genetic information (genotypes) and visible traits (phenotypes). GS predicts desired genetic breeding values for breeding and tree improvement populations. In GS applications, algorithms are used to rank and select trees based on one or more traits of interest simultaneously.

GS offers several advantages for second and third generation breeding programs, including shortened breeding cycles with reduced reliance on field testing, improved evaluation of quantitative traits, improved traceability for pedigree, and accelerated delivery of genetic gain while maintaining genetic diversity. If widely adopted, GS has the potential to be a game-changer in tree-breeding that enables the use of a range of data (e.g. pedigree, phenotype, genetic, environmental variation, climate variables, and so forth) to inform breeding selections.

As with all novel technologies, uncertainty is inherent in the application of GS. Assessment and communication of uncertainties are important features of decision-making. This document draws on best-practice guidelines...

- communicating uncertainty in accessible language,
- acknowledging multiple viewpoints and values,
- promoting transparent justifications for decisions.[iii]

Key Takeaway Messages

Assessment and communication of uncertainty is important for forest management, particularly in face of novel technologies and climate change.

Numerical expressions of uncertainty alone are not sufficient.

Deeper forms of uncertainty – such as ambiguity and ignorance - are harder to assess and communicate.

Multiple perspectives and values are important to take into consideration.

Genomic Selection for Forestry

Canada's forests are facing unprecedented challenges, including climate change, drought, pests, and pathogens. To ensure sustainability of Canada's forest industry, selecting, breeding, and planting trees that are resistant to such stresses is of utmost importance.[iv]

GS is a biotechnology application designed to help tree breeding operations become more efficient and effective. While GS is relatively new in forestry, it was developed for and has been successfully applied in agriculture, with the most advanced use to date in dairy cattle. GS is also used in breeding programs for beef cattle, chickens, pigs, and a variety of major agricultural crops. GS is now in initial stages of application for forest breeding operations across Canada, the United States, Europe, and Latin America.

Most traits of interest for forestry are complex and involve the interaction of many genes scattered throughout the entire DNA of an organism, referred to as the genome. Advanced sequencing technologies and powerful computational algorithms enable scientists to identify genetic markers that are closely linked to genes that contribute to the expression of phenotypic traits.

Conifer trees have been on Earth for hundreds of millions of years. Conifer species have extremely large genomes, with low variation in number of chromosomes. For example, most members of the pine family have 12 very large chromosomes, compared with humans who have 23 chromosomes with significantly smaller amounts of DNA. Despite their economic and ecological importance, conifer species are understudied at the genetic level. This situation is changing with the development of next generation sequencing technologies. The genomes of Norway spruce, white spruce, and loblolly pine have already been sequenced, with more to come.[v]

DNA is coiled into a shape called a double-helix molecule made up of four bases (Adenine, Thymine, Cytosine, Guanine), where A joins with T, and C joins with G. Particular sequences of these nucleic acids code for amino acids which build the proteins that make up all living things.

Genomic selection involves examining genomes for variants such as SNPs: single nucleotide polymorphisms. A SNP is a place in the genome where one nucleotide differs between individuals. While many SNPs have no effect on phenotype, some may affect observable traits directly or, more frequently, be closely associated with nearby regions of DNA that affect a given trait.

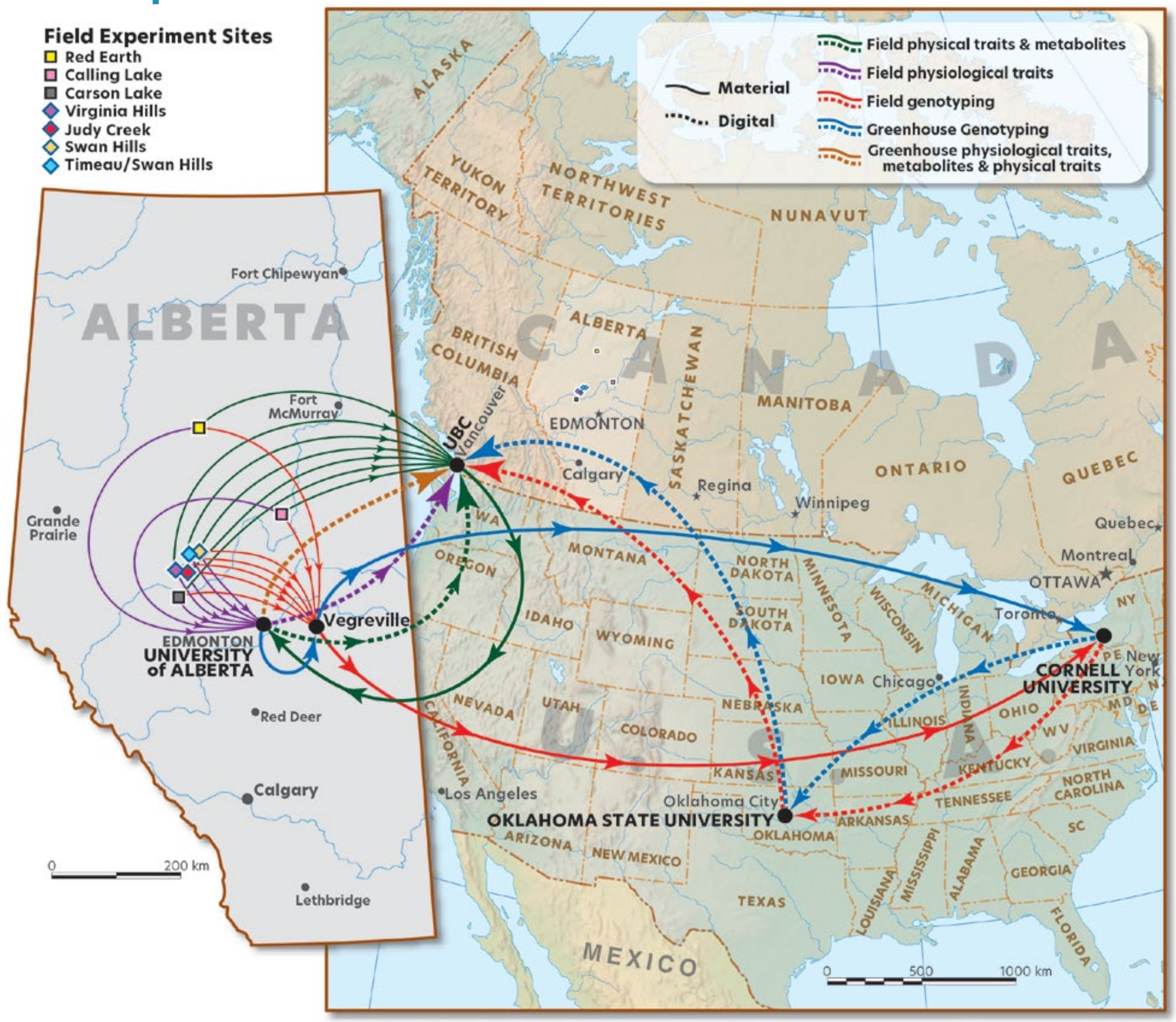
The need for uncertainty communication is recognized for GS applications in forensic science and medicine.

Background

GS was first proposed in the early 2000s. At that time, the technological capacity to sequence genomes was not available.[vi] This changed with the completion of the Human Genome Project in 2003. Cost-effective, high-throughput sequencing technologies (called next generation sequencing technologies) emerged thereafter, enabling GS to move from concept to application in fields as diverse as medicine, agriculture, and forestry. The genetic sequencing of human, animal, and plant DNA continues to increase exponentially. The need for uncertainty communication is recognized for GS applications in forensic science and medicine.

How is a genomic selection model developed?

The development of GS is a significant undertaking that involves international collaborations among scientists and research institutes. The complexity of research and the flows of information across institutions and borders is one of several sources of uncertainty. The following diagram shows the process of data collection and analysis in the RES-FOR project.



Source: Myles, 2020[vii]

Data gathering

Tree needles (for DNA extraction) and wood samples gathered from tree improvement populations. Existing (30 year +) height and diameter data and pedigree information also used for models.

Extraction & Sequencing

Tree needles sent to Vegreville for DNA extraction, and then Cornell for DNA sequencing. Technicians analyze DNA with a sequencing machine that breaks DNA into smaller pieces. The sequence of bases is then read, one base at a time.

Analysis

A team of experts at Oklahoma State University and UBC use specialized software and powerful computers to piece together and compare DNA fragments for analysis of variants. Single point variations are called SNPs (single nucleotide polymorphisms). Statistical algorithms find patterns in genomic and phenotypic data to create models.

Genomic Estimated Breeding Value

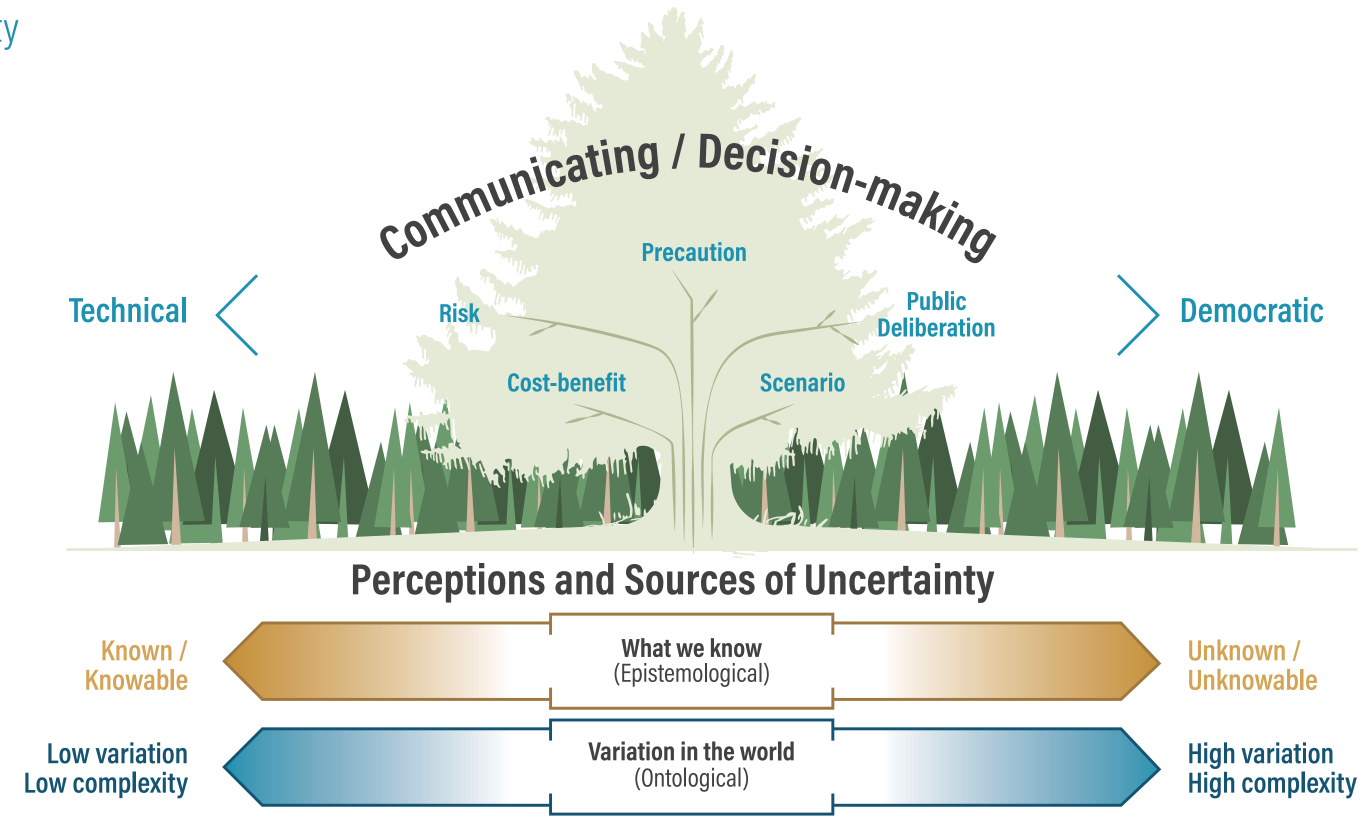
Shiny R application used to visualize the results from different selection methods.

Defining, assessing and communicating uncertainty

Uncertainty refers to limitations in knowledge, particularly in the face of complexity and variability. Uncertainty is not the same thing as risk, which implies prior knowledge of the possible likelihood and magnitude of outcomes, and is generally used to assess the likelihood of harm.[viii] Uncertainty assessment extends far beyond risk assessment: it is a means of informing decision-makers, users and publics of multiple factors that might limit our ability to predict specific outcomes during application.

Uncertainty assessment involves:

1. Acknowledging perceptions and sources of uncertainty
2. Integrating uncertainty into communication and decision-making



Perceptions of Uncertainty

Uncertainty is understood and experienced in different ways by different people. A person's perceptions of and responses to uncertainty are shaped by emotions, values, and experiences, which lend themselves to highlighting some sources of uncertainty as more important than others.

Depending on the technology or application in question, different groups, representing multiple perspectives, may express an interest. Certain new technologies may only have relevance to a small and relatively homogenous group, whereas other technologies might be of interest to more people.

Feelings. In response to uncertainty, people can experience a range of feelings, including excitement and anxiety, to differing degrees depending on the individual, and the stakes involved.

Values. What we value informs how we react to uncertainty. In turn, facts and values often merge under conditions of uncertainty. For example, 'uncertainty' can evoke unease at the loss of control over one's environment, or can foster doubt, skepticism and conflict.

Experiences. Previous lived experiences, negative or positive, may prime individuals to be sensitive to certain potential outcomes, leading either to over-confidence or heightened concern, about new technologies.

Evaluate Perceptions of Uncertainty

No one is immune to feelings, values, and experiences when it comes to assessing uncertainty. The following questions are important to ask to better understand our personal responses to uncertainty.

What concepts come to mind when you hear the term uncertainty?

What feelings arise when you think about uncertainty?

What strategies do you typically use to cope with uncertainty?

How has your professional training influenced how you approach uncertainty?

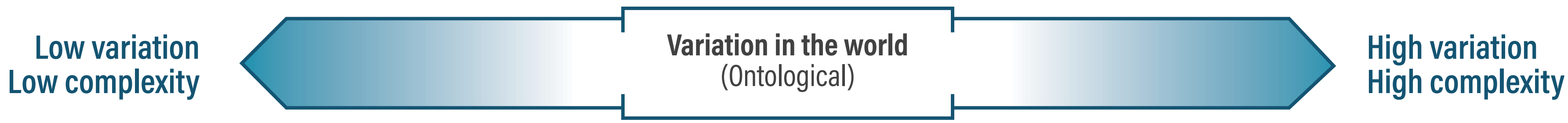
Sources of Uncertainty

The sources of uncertainty are numerous, but generally fall into two categories:

Variation in the world (**ontological uncertainty**), and Human limits of knowledge (**epistemological uncertainty**).

Variation in the world

Nature and society are complex systems, characterised by the dynamic interaction of multiple factors and processes. In complex systems, outcomes emerge from these interactions, and are often difficult or impossible to predict. Sources of variability that come into play with the application of novel technologies are broad-ranging, and include natural, technical, cultural, and socio-political systems. The messiest domain is often the socio-political system, which is also most excluded from consideration in expert-led uncertainty assessments.



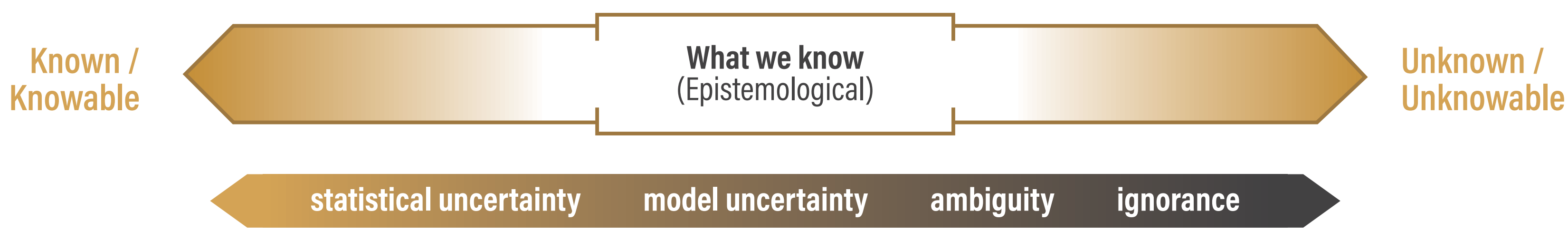
Limitations of knowledge: What we know

Statistical uncertainty involves quantitative procedures for assessing uncertainties related to measurement and estimated variation. Statistics are also used to make associations between genotypes and phenotypes through regression analysis, and to impute ('fill in the blanks') where data are missing.

Model uncertainty refers to reliability of models. Models are simplifications of reality in which complexity is intentionally left out to focus on variables of interest. As statistician George Box famously remarked, "all models are wrong, but some are useful." [ix] The point that Box was making is that all models are representations of reality and contain uncertainties. While some model uncertainties can be expressed with quantitative values, uncertainties also arise from assumptions and other unknowns that cannot be expressed meaningfully through numbers. The extent to which model uncertainty can be reduced is a point of debate.

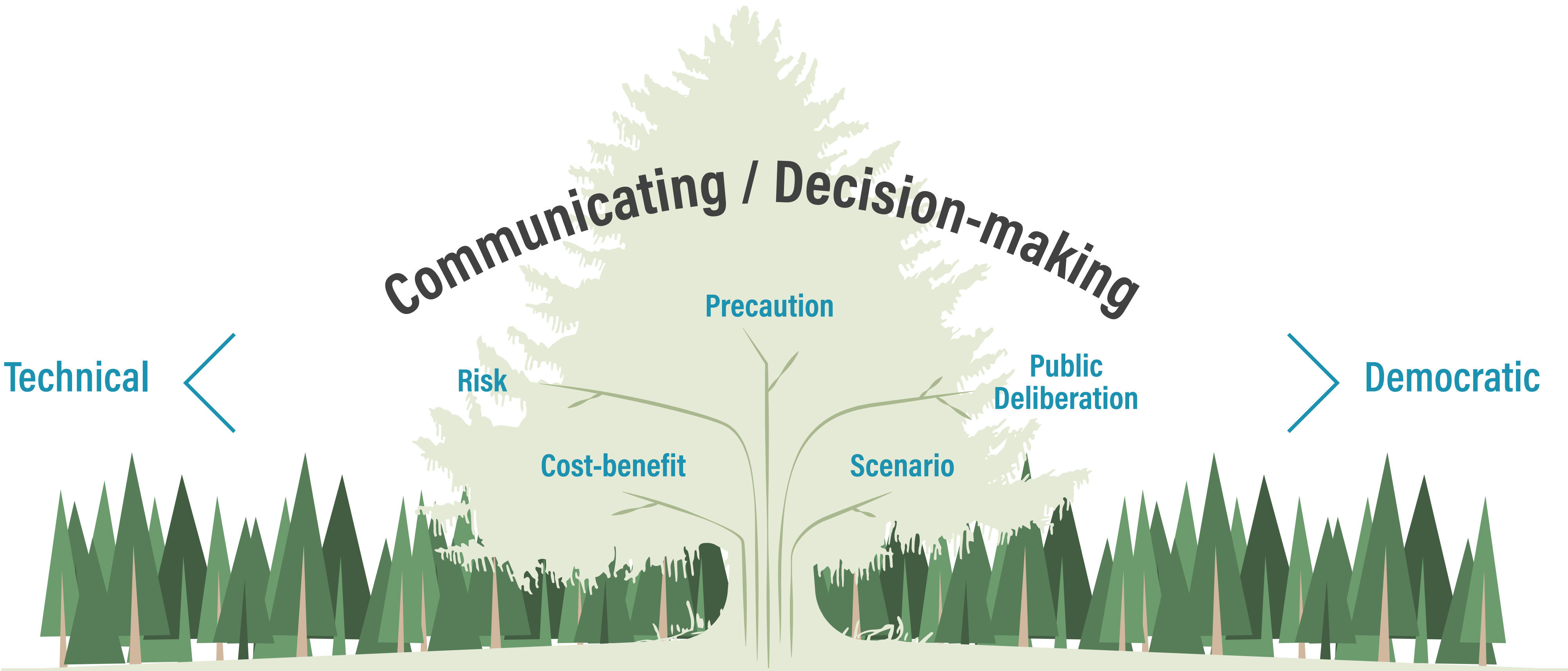
Ambiguity refers to different worldviews, perspectives, and values. Ambiguity can arise from different interpretations of research findings, and also of potential harms and benefits. If ambiguity is not identified, conflict can arise as important perspectives may be overlooked.

Ignorance refers to situations in which "we don't know what we don't know". Ignorance can refer to limitations of both individual and communal knowledge.[x] Acknowledgement of ignorance can temper overconfidence in expected outcomes.



Communication and decision-making

Various decision and communication tools are available to address uncertainty, some are more widely deployed than others. These approaches are not necessarily mutually exclusive, and can be combined depending on the purpose and context.[xi]



Risk assessment involves identifying the likelihood of negative outcomes based on known estimates of the probability and magnitude of events.

Cost-benefit analysis (CBA) shares many elements with risk assessment, applied specifically to economic uncertainty. CBA typically employs efficiency as a core value.

Precautionary approaches are used if available evidence suggests the potential for a novel technology to pose significant harm, but is incomplete or insufficient to inform risk analysis. Precautionary approaches call for a broader approach to assessing uncertainty, and often engage multiple viewpoints and values.

Qualitative scenario analysis involves presenting and comparing various plausible visions of the future. Scenario analysis is particularly useful when high levels of uncertainty have been acknowledged.

Public deliberation refers to structured discussions that involve diverse values, interests, and perspectives. Public deliberation is especially warranted in conditions of high uncertainty and ambiguity. Key values are transparency, inclusivity, and diversity. Reflexivity about values, humility about the limitations of knowledge, and openness to multiple ways of knowing and valuing the world are essential for ensuring that public deliberation enhances decision-making.

	Risk Assessment	Cost Benefit Scenarios	Precautionary Approaches	Qualitative Scenarios	Public Deliberation
Involves diverse values	✗	✗	✓	✓	✓
Requires technical expertise	✓	✓	✓	✗	✗
Addresses knowns	✓	✓	✓	✓	✓
Addresses unknowns	✗	✗	✓	✓	✓

Genomic Selection in Forestry
Potential types of Uncertainty

Novel technologies such as GS will introduce and reduce uncertainties, which is the focus of this document. However, assessments of uncertainty also need to take into account the uncertainties that can arise in conventional approaches to tree selection that do not employ GS.

The following table outlines some examples of potential types of uncertainty associated with the introduction of GS for forestry. The intent is to raise questions for further discussion. Please note that not all of the questions have immediate answers.

Potential sources of uncertainty	Examples	Questions
Nature	▪ **Climate change** will contribute to unpredictable and variable seasonal weather with different impacts across growing seasons. As temperatures increase, trees will face more pressure from drought, disease and insects. Yet, regional impacts of climate change are difficult to anticipate and different models give different results ▪ **Forests** and **conifers** are complex systems. Conifers have large and relatively poorly-understood genomes	▪ How effective are conventional tree breeding methods in dealing with potential impacts of climate change? ▪ How effective will GS be in dealing with potential impacts of climate change? ▪ What assumptions about climate change are made in breeding selections (with or without GS)? ▪ What if conifers cannot adapt to climate change on their own?
Technology	▪ **Statistical methods** are diverse and can give rise to different outcomes ▪ Requirements for **database storage and management** may change as GS technologies evolve ▪ **New technologies** require new forms of expertise and training	▪ What are the implications of using GS for tree improvement? ▪ What are the implications of not using GS for tree improvement? ▪ Will GS make selections that produce a healthy forest in the long term? ▪ Will conventional methods make selections that produce a healthy forest in the long term?
Values	▪ The goals, values and concerns of users, regulators and publics may shift over time ▪ Key values associated with GS (optimization, efficiency, improvement) may not be held by other stakeholders	▪ How will the costs and benefits of GS be distributed, in the short and long term? ▪ What is valued in conventional tree breeding selections? How might GS alter these values?
Socio-Political	▪ **Economic markets** are volatile, particularly for forest products. ▪ Changes in **forest management policies** may impact GS ▪ Shifting terrain of **environmental politics** may affect public support, including Indigenous rights	▪ How will GS influence forest management decisions under existing policy regimes (e.g. Annual Cut Effect)? ▪ Are different policy frameworks necessary to support the uptake of GS? ▪ To what extent are local communities and Indigenous groups interested in GS applications for forestry?

Recommendations for Communication and Decision-making

A **risk management** approach is the dominant framework for biotechnology in Canada. Considering the brief uncertainty assessment provided, a risk management approach might overlook multiple sources of uncertainty in ways that exclude input from various stakeholders. Risk management often prescribes a one-way, 'educate the public' type of communication that can lead to conflict.

A **precautionary approach** may not be warranted, because at this time, the potential for significant harm has not been identified scientifically. Precautionary approaches can lead to restrictions in implementation that prevent further innovation and learning.

Scenario-based approaches may be applicable to GS applications in forestry. For example, in the RES-FOR project, the results from the multi-objective optimization algorithms used to calculate genomic breeding values are communicated to end-users through the Shiny R application. This allows users to select, weigh, and compare traits. Qualitative scenarios – presented as plausible storylines depicting possible outcomes – can enhance participation with non-technical audiences.

Public deliberation is particularly important in situations in which multiple perspectives and values are at play. While public deliberation requires sustained funding, the cost is nominal compared with the amount of public money already invested in GS research in Canada. In the RES-FOR project, public deliberation was enhanced through highlight sheets, briefing documents, and workshops with users and decision-makers.

References

i Saltelli, A et al. 2020. Five ways to ensure that models serve society: A manifesto. *Nature*. June 24. Available at: https://www.nature.com/articles/d41586-020-01812-9; Blastland et al, 2020. Five rules for evidence communication. ***Nature***. Nov. 28. https://www.nature.com/articles/d41586-020-03189-1

ii Blue G, Davidson D. 2021. Co-producing uncertainty in public science: The case of genomic selection in forestry. Public Understanding of Science. https://doi.org/10.1177/0963662520982540

iii Saltelli, A et al. 2020.

iv Naidoo, S, Slippers, B, Plett, J. Coles, D. Oates, C. 2019. The Road to Resistance in Forest Trees. *Frontiers in Plant Science* 10 (273). doi: 10.3389/fpls.2019.00273

v Prunier, J., Verta, JP., MacKay, J. 2015. Conifer genomics and adaptation: At the crossroads of genetic diversity and genome function. New Phytologist 209 (1): 44 – 62.

vi Meuwissen, T.H., Hayes, B.J. & Goddard, M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics, 157, 1819-29.

vii Myles, K. 2020. Opening-up to broaden out: A case study of representational practices as inputs for social assessments of genomic selection technologies for forestry. (Unpublished master's thesis). University of Calgary, Calgary, Alberta.

viii Knight, Frank. 1921. *Risk, uncertainty, and profit.* . Boston: Houghton and Mifflin.

ix Box, George 1976. Science and statistics. Journal of American Statistical Association 71:791–799. doi: https://doi.org/10.1080/01621459.1976.10480949.

x Firestein, S. 2012. Ignorance: How it drives science. Oxford, UK: Oxford University Press.

xi Committee on Decision Making Under Uncertainty; Board on Population Health and Public Health Practice; Institute of Medicine. 2013. Environmental Decisions in the Face of Uncertainty. Washington (DC): National Academies Press (US); May 20. Available from: https://www.ncbi.nlm.nih.gov/books/NBK200848/ doi: 10.17226/12568; Scoones, I., Stirling, A. 2020. The politics of uncertainty: Challenges of transformation. London and New York: Routledge.

Acknowledgements

We acknowledge cash funding for this research from Genome Canada, Genome Alberta through Alberta Economic Trade and Development, Genome British Columbia, the University of Alberta and the University of Calgary. Further cash funding has been provided by Alberta Innovates BioSolutions, Forest Resource Improvement Association of Alberta, and the Forest Resource Improvement Program through West Fraser Ltd. (Blue Ridge Lumber and Hinton Wood Products) and Weyerhaeuser Timberlands (Grande Prairie and Pembina). In-kind funding has been provided by Alberta Agriculture and Forestry, Blue Ridge Lumber West Fraser, Weyerhaeuser Timberlands Grande Prairie, and the Thomas, Wishart and Erbilgin labs in support of the Resilient Forests (RES-FOR): Climate, Pests & Policy – Genomic Applications project.

RES-FOR

For more information on the RES-FOR project please contact:

Barb Thomas: bthomas@ualberta.ca or Stacy Bergheim: sberghei@ualberta.ca

Resilient Forests: https://resfor.ualberta.ca/

For more information on this document, please contact:

Debra Davidson: ddavidso@ualberta.ca or Gwendolyn Blue: ggblue@ucalgary.ca