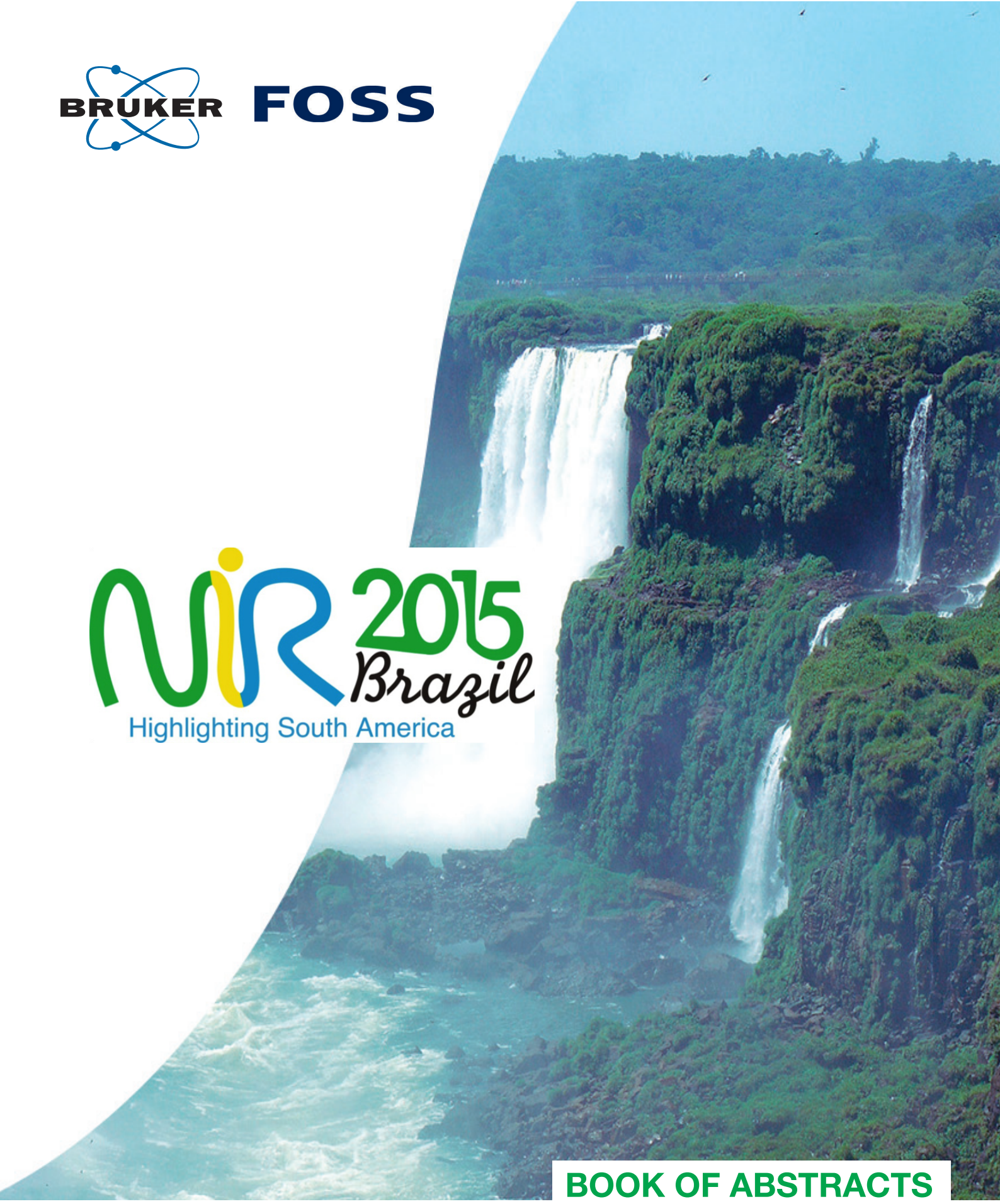
Determination of Larch Taxa with Near Infrared Spectroscopy

Vincent Segura, Jean-Paul J.-P. Charpentier, Kévin Ader, Corinne Buret, Nathalie Boizot, Luc Eric Pâques

BRUKER FOSS



NIR 2015 Brazil

Highlighting South America

BOOK OF ABSTRACTS

OCTOBER
18-23

17th INTERNATIONAL CONFERENCE
ON NEAR INFRARED SPECTROSCOPY

FOZ DO IGUASSU BRAZIL

Determination of Larch Taxa with Near Infrared Spectroscopy

Vincent Segura, Jean-Paul Charpentier, Kévin Ader, Corinne Buret, Nathalie Boizot, Luc Eric Pâques

INRA, UR0588 Amélioration, Génétique et Physiologie Forestières

Topic: Wild Live/Ecology/Environmental

In the context of forest tree breeding, a strong quality control is required especially when the improved varieties derive from interspecific hybridization. This is the case of Larch whose improved varieties are hybrids between European and Japanese species. In this case, the quality control faces two challenges: (1) verifying the parental species in seed orchards, and (2) attesting the hybrid status of seedlings in nurseries. In both cases, a cost effective high-throughput technique is required to identify the genetic origin of the trees. In this context, near infrared spectroscopy (NIRS) is a potentially interesting alternative to more traditional barcoding approaches based on morphological or molecular markers which are typically tedious and/or costly. The use of NIRS for barcoding relies on the assumption that some chemical and/or physical properties are able to discriminate the genetic origin of the analyzed samples (Cruickshank and Munck, 2011). A recent study in Pine has successfully tested this hypothesis for leaf samples (Meder et al., 2014), underlining the potential of NIRS for quality control in forest tree breeding programs. The present study aims at evaluating this approach on a total of 350 leaf samples that have been collected over three consecutive years on European, Japanese and hybrid Larches that grow on the same site. We first performed a principal component analysis of the spectral dataset which revealed some grouping according to the species and sampling year, suggesting that NIRS is able to capture both useful (genetic) and undesirable (year) information. In addition, the ability of NIR spectra to capture the useful information varied depending on the statistical pretreatment of the spectra. Then, we assessed the ability of NIRS to predict the genetic origin of our samples with a partial least square discriminant analysis. Calibrations were first carried out by year of sampling and the predictive ability of the models was assessed through a repeated cross-validation. The corresponding models included between 4 and 14 latent variables, and had a prediction accuracy which ranged between 0.80 and 0.95, depending on the year of harvest. The robustness of these models was further evaluated by external validations across the sampling years. As a consequence of the large between-year variation observed in the NIRS dataset, the prediction accuracy decreased across years with values ranging between 0.64 and 0.89. We thus tried to obtain a more robust model trained over the three harvest years with two third of the data and validated with the remaining third. The resulting cross-validation and validation accuracies were in the same range, close to 0.9, underlining the usefulness of NIRS for determining the genetic origin of larch samples. Interestingly, we found in the repeated cross-validation schemes that several samples were systematically wrongly predicted, suggesting some mistakes in the genetic origin of these samples. Such mistakes have likely lead to a downward bias in our prediction accuracy estimates. Some molecular analyses using DNA barcoding are thus currently under progress in order to test this hypothesis and update our prediction accuracy.

Acknowledgments

We are grateful to the French ministry of food agriculture and forest for funding the study through the program “DGPAAT Evaluation des variétés forestières”.