Alternative approaches to bioclimate envelope mapping that could be leveraged to update Climate-FVS suitability scores and other future projections of climatic suitability for tree species.

Prepared by David Diaz and Matt Perry (Ecotrust)

Several research studies have shown that purely statistical approaches such as Random Forests are commonly observed to over-predict climatic vulnerability and species' range losses in comparison to process-based models in response to projected climate changes (e.g., Morin & Thuiller 2009; Keenan *et al.* 2011; Cheaib *et al.* 2012).

In this exploratory analysis, we have closely followed the Random Forests approach used by Crookston et al. (2010) that currently drives the calculation of climatic suitability scores that are factored into Climate-FVS.

A good example this over-prediction issue for Random Forests can be seen in following map set, particularly the prediction of current climatic suitability for Douglas-fir when all FIA presence/absence data is used as training data (bottom row of map series). Although it is well-known that Douglas-fir is capable of growing in the Willamette Valley, the occurrence of FIA plots with Douglas-fir absent in the Willamette Valley (see bottom-right map), as well as the paucity of other FIA plots in the area contribute to a machine learning response that the climate is not suitable for this species in this location. We believe some of this spatial over-fitting is likely due to excessive influence of 'false negatives.' For example, the absence of Douglas-fir in FIA plots in the Willamette Valley is likely due to human land-use, competition with other species, and similar factors that are ***not*** related to the climatic suitability for the species to grow. If these FIA absences are used in the training data, if/when future climatic conditions in other locations begin to resemble those of the Willamette Valley (or other places with 'false negatives'), the regression model will predict unsuitability for the species.

We have started evaluating alternative methods to curtail the influence of 'false negatives' in the Random Forests approach by adjusting the training data sets used. Two of these approaches are shown in the top 2 rows of maps. The basic approach is using a species distribution map predicted by a process-model such as 3-PG (see top row of maps) or from expert judgment such as the Little Species Distribution maps to exclude FIA plot absences from the training data set in the spatial extents that these methods consider suitable for th e species. The underlying logic is that the occurrence of FIA plots within these zones where Douglas-fir is not observed can be thrown out of the training dataset as 'false negatives.' Thus, as the Random Forests regression models begins to learn the relationships between climatic conditions that are unsuitable for Douglas-fir and FIA plot records, it will only be considering FIA plots where Douglas-fir is absent outside the species' range.

The current climatic suitability can be seen to much more closely resemble the known species range in this approach, and is likely to offer a much more reliable mapping of future climatic suitability under climate change. There is a direct opportunity to integrate an approach like this into the suitability scoring system that drives Climate-FVS, but also more broadly to consider this approach for integrating process-models and expert judgment into more refined bioclimate envelope projections in general.

Cheaib, A., Badeau, V., Boe, J., Chuine, I., Delire, C., Dufrêne, E., François, C., Gritti, E.S., Legay, M., Pagé, C., Thuiller, W., Viovy, N. & Leadley, P. (2012) Climate change impacts on tree ranges: model intercomparison facilitates understanding and quantification of uncertainty: Understanding and quantification of uncertainties of climate change impacts on tree range. *Ecology Letters*, **15**, 533–544.

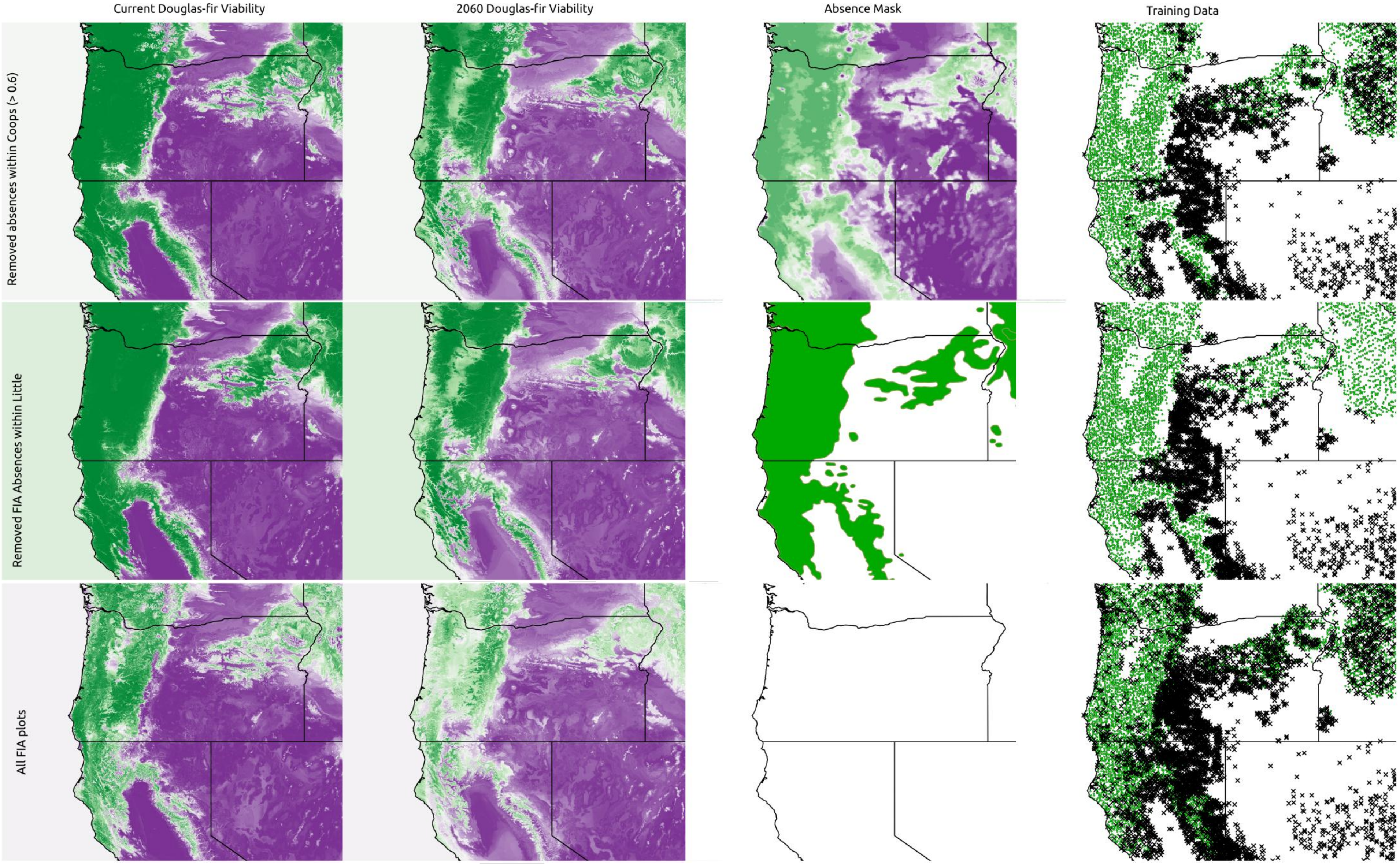
Coops, N.C., Waring, R.H., Beier, C., Roy-Jauvin, R. & Wang, T. (2011) Modeling the occurrence of 15 coniferous tree species throughout the Pacific Northwest of North America using a hybrid approach of a generic process-based growth model and decision tree analysis: Modeling the occurrence of fifteen coniferous tree species. *Applied Vegetation Science*, **14**, 402–414.

Coops, N.C., Waring, R.H. & Schroeder, T.A. (2009) Combining a generic process-based productivity model and a statistical classification method to predict the presence and absence of tree species in the Pacific Northwest, U.S.A. *Ecological Modelling*, **220**, 1787–1796.

Crookston, N.L., Rehfeldt, G.E., Dixon, G.E. & Weiskittel, A.R. (2010) Addressing climate change in the forest vegetation simulator to assess impacts on landscape forest dynamics. *Forest Ecology and Management*, **260**, 1198–1211.

Keenan, T., Maria Serra, J., Lloret, F., Ninyerola, M. & Sabate, S. (2011) Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO2 matters! *Global Change Biology*, **17**, 565–579.

Morin, X. & Thuiller, W. (2009) Comparing niche-and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology*, **90**, 1301–1313.



These maps show Random Forests predicted scores for current and future climatic suitability for Douglas-fir based on alternative approaches to filter which training data (FIA presence/absence) inform climate-species regression relationships from the 1961-1990 climate baseline used by Crookston et al. (2010) to create suitability scores for use in Climate-FVS. The bottom row closely mimics the approach that is currently in place for Climate-FVS, and considers all presence/absence data from available FIA plots. The middle row uses the Little Species Distribution for Douglas-fir to remove FIA 'absence' data within the species' range from the training data. The top row is a similar approach using a process-model approach (3PG) employed by Coops et al. (2009; 2011) *to mask out FIA plot absences within the range identified as currently climatically suitable for Douglas-fir.*