Photo: © elPadawan, Flickr

“The future is already here — it's just not very evenly distributed”, William Gibson

Very primitive statistical models
(only phenotypic data)

Advanced statistical models
(phenotypic+genomic data)

Compromises in

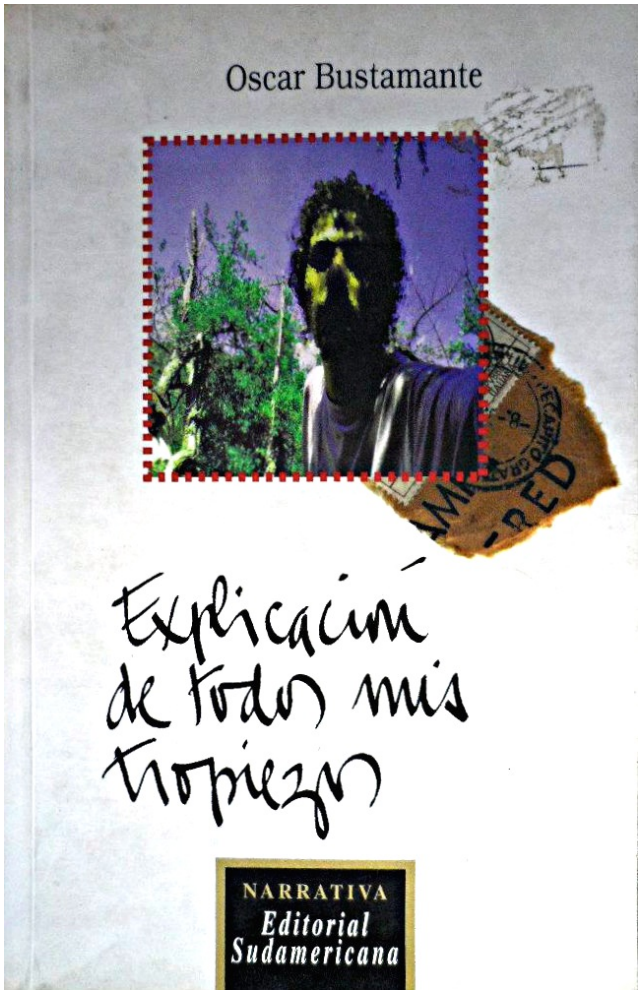
The genetic evaluation system of radiata pine in New Zealand

Luis A. Apiolaza

School of Forestry, University of Canterbury, Christchurch, New Zealand

Luis.Apiolaza@canterbury.ac.nz — http://apiolaza.net

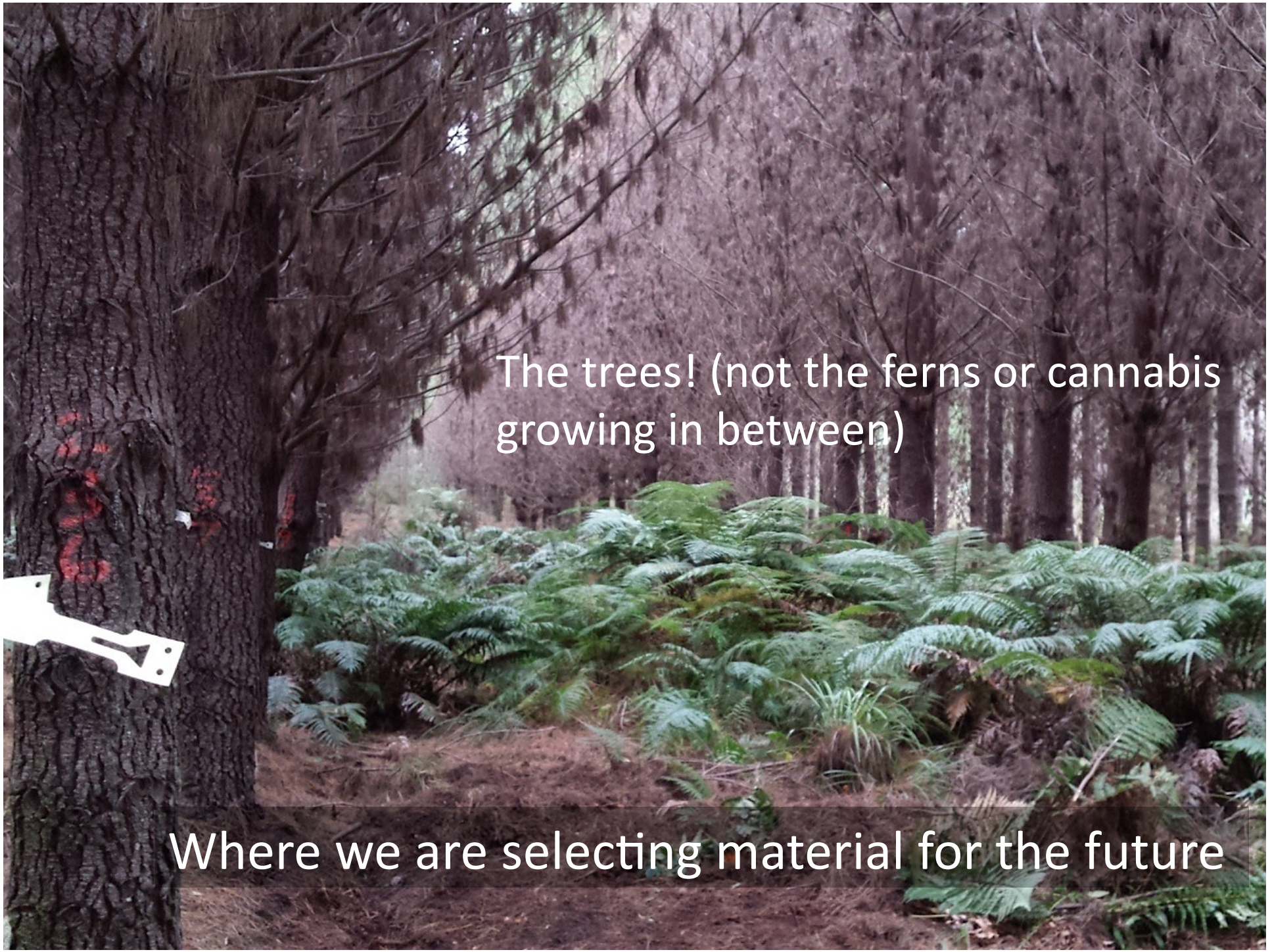
UC
UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND



"Explanation of all my difficulties"


Where we come from (1944)


Where we are now (2014)
Photo: © Horst Kiechle, Flickr

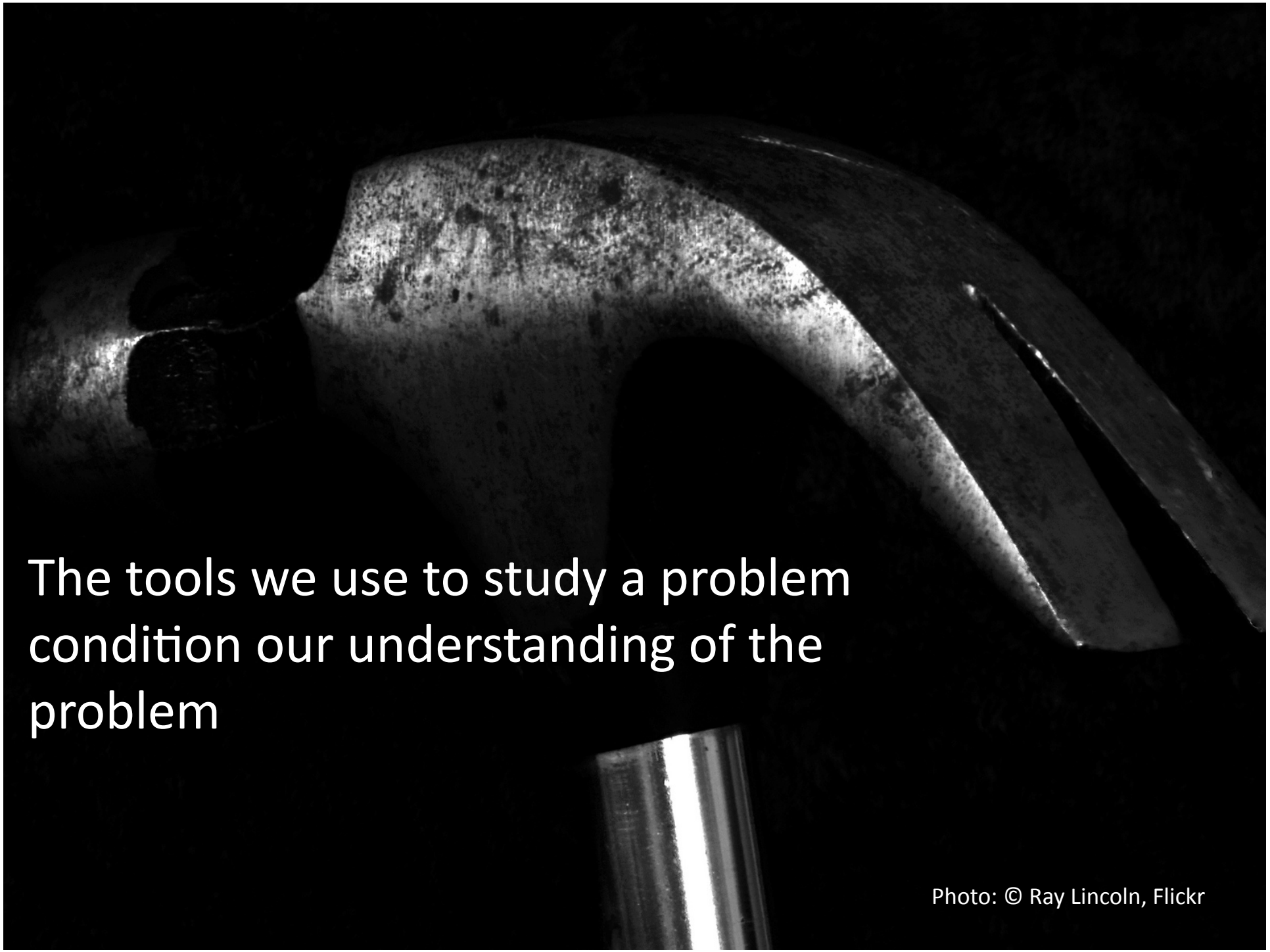

The trees! (not the ferns or cannabis growing in between)
Where we are selecting material for the future

Clearly there has been progress

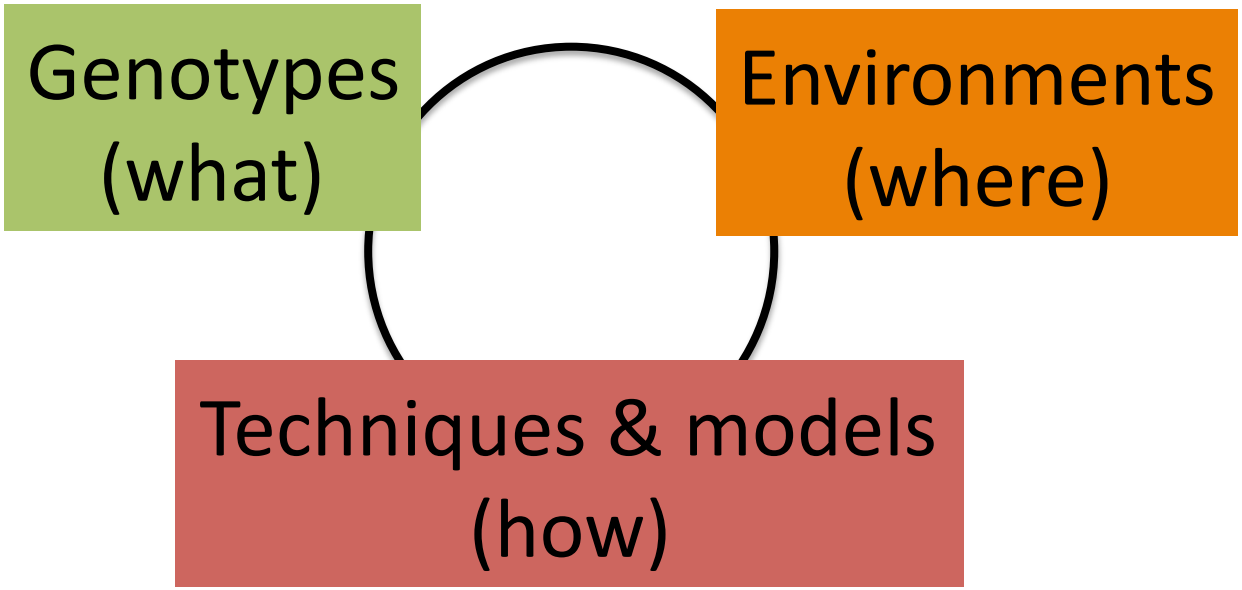
but

Has it been enough for 50 years?

How far from our potential?


The tools we use to study a problem condition our understanding of the problem
Photo: © Ray Lincoln, Flickr

As breeders, our **genetic evaluation system** is the main tool we use to understand the world



Have an effect on our conclusions

Genotypes (what)

3,000 parents under testing

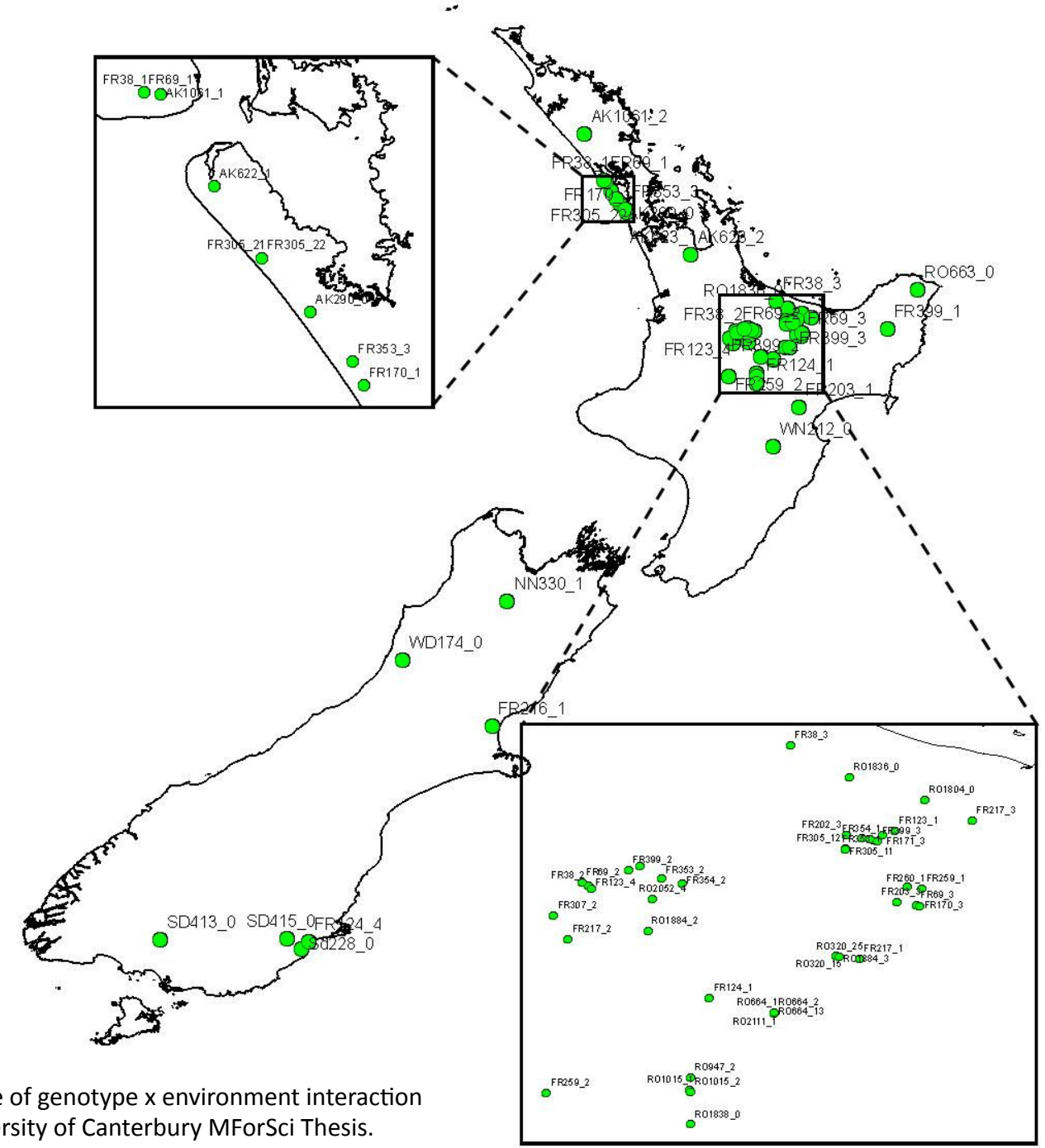
325,000+ progeny in trials

Multiple experimental designs

Multiple mating designs (OP, CP, clonal)

Where
(with data already)

82 trials



Timothy McDonald 2009 Making sense of genotype x environment interaction of *Pinus radiata* in New Zealand. University of Canterbury MForSci Thesis.

Techniques & models (how)

Two extreme naive models for genetic evaluation:

Univariate analysis, homogeneous variance, equal correlation between all sites

Generic understanding of GxE interaction

Ideally our model would be in between

Multivariate analysis, unstructured heterogeneous variances, all sites highly interacting with different correlations

Every site is highly interacting

Techniques & models (how)

Plant breeding

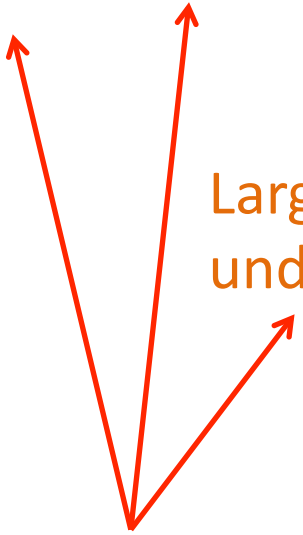
Emphasis on experimental design (including spatial trends)

Few genotypes under testing

Animal breeding

Emphasis on pedigree

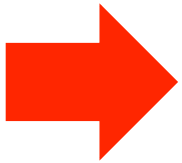
Large number of genotypes under testing

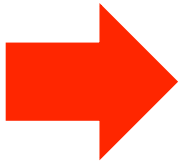


Tree breeding is special

This creates 2 big problems in tree breeding

(Remember that gold standard is single-stage evaluations)

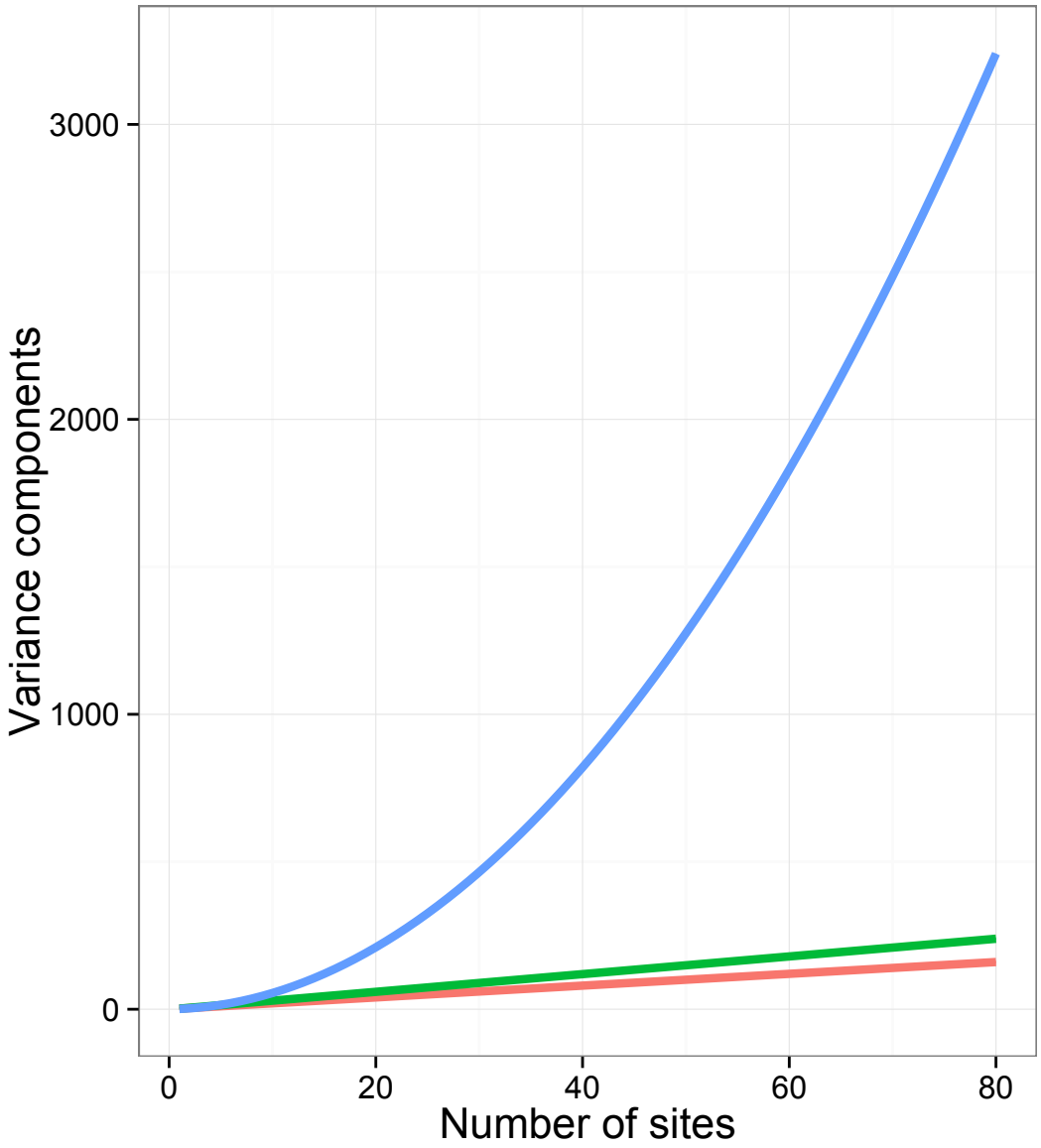
We have too many traits  Factor analytic models

We have too many genotypes  Reduced animal models

We use both **simultaneously** in the New Zealand Radiata Pine evaluation.

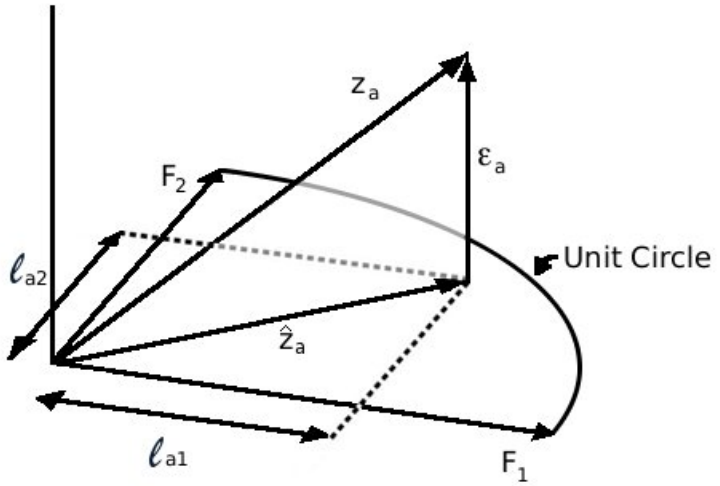
Therefore we can run a single-stage evaluation

Number of **G** parameters to estimate



Projecting sites on 1 (FA1) or 2 (FA2) latent variables:

1) Is often enough to represent the correlation structure for 80 sites.
2) Reduces estimation problems



https://en.wikipedia.org/wiki/Factor_analysis

Reduced Animal Model*

Makes a distinction between trees with progeny (roughly 3,000 in our problem) and trees without progeny (>300,000), greatly reducing the size of the problem.

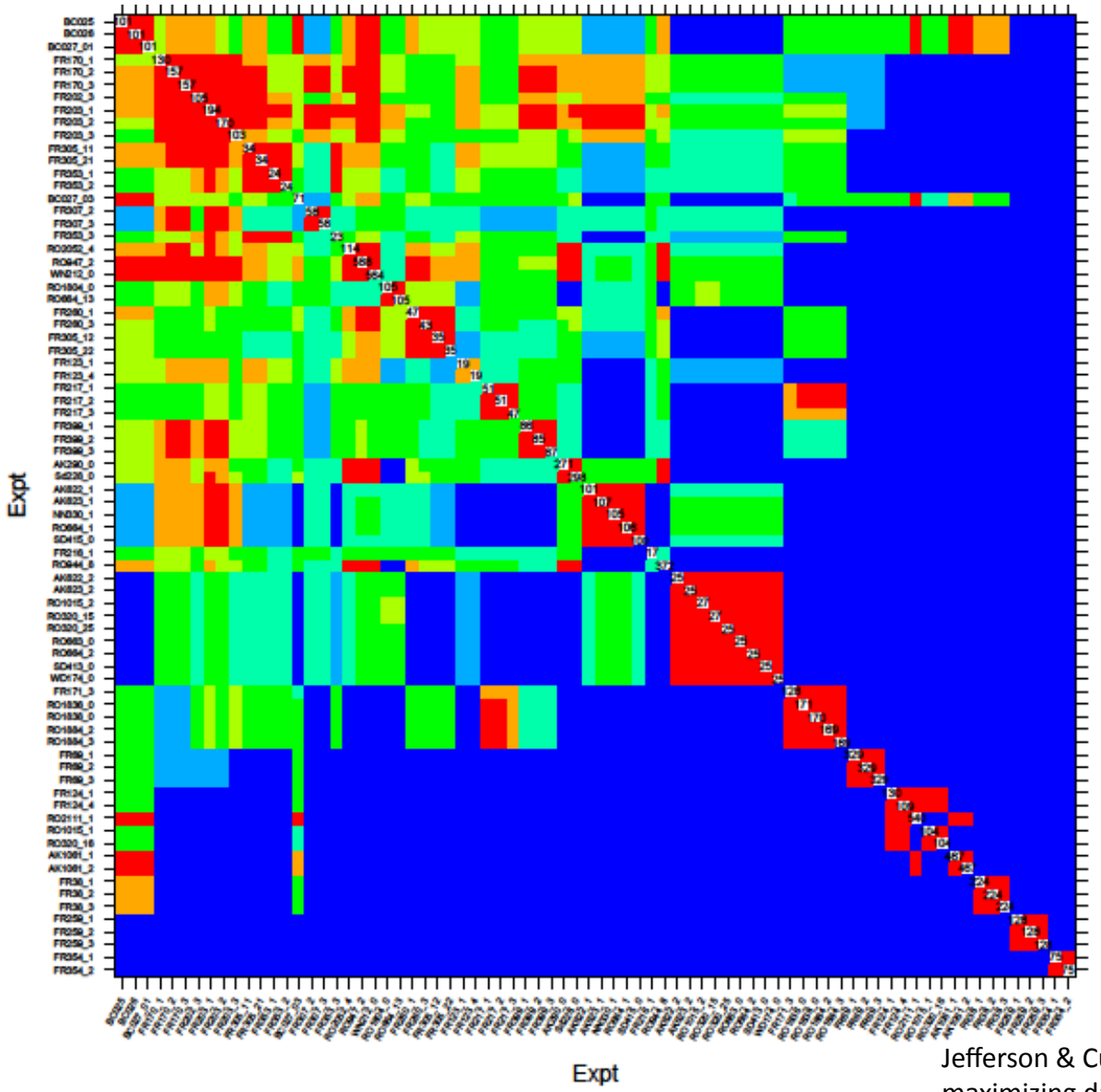
$\mathbf{A}_{pp}^{-1}$ is diagonal for parents, ignores Mendelian sampling for non-parents, produces breeding values only for parents. If required, model can be modified to obtain forward selections, by expanding $\mathbf{A}_{pp}$

On top of that, the NZRPBC uses a Factor Analytic structure to model the reduced animal model**.

*Pollak and Quaas (1980)

**Cullis, Smith, Jefferson & Thompson 2013. Implementation strategy for RPBC breeding values incorporating GxE

(dis)connectedness



Parents in common

0
5
10
15
20
20+

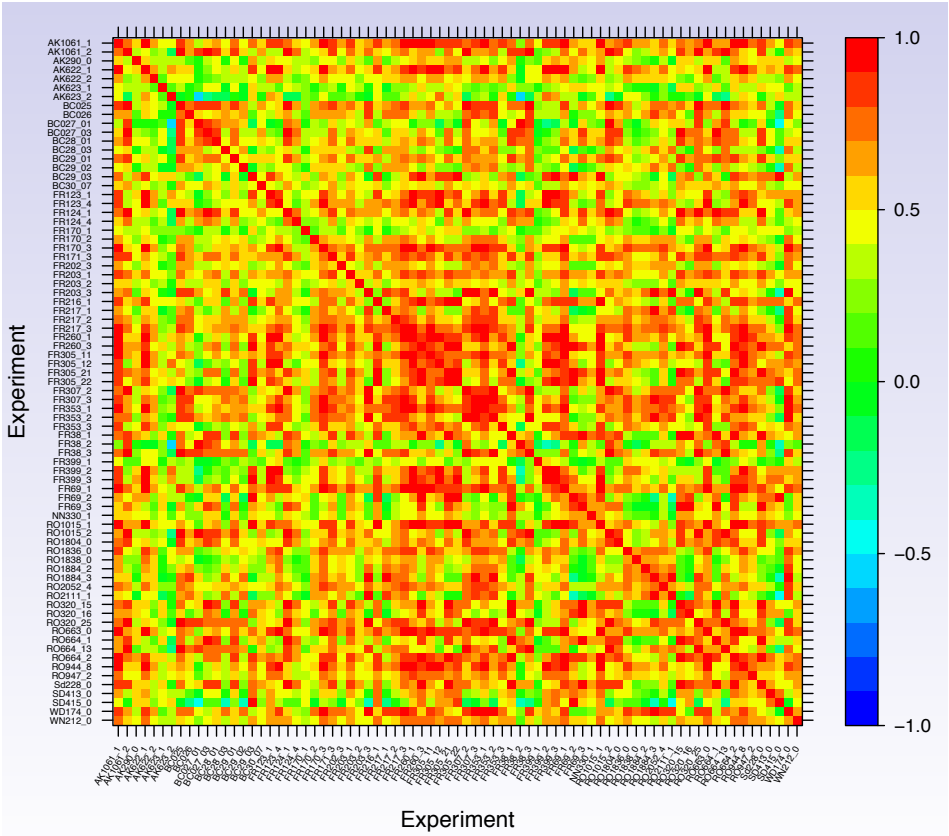
Jefferson & Cullis 2012. Prediction of breeding values maximizing data from trials over 76 sites.

(dis)connectedness

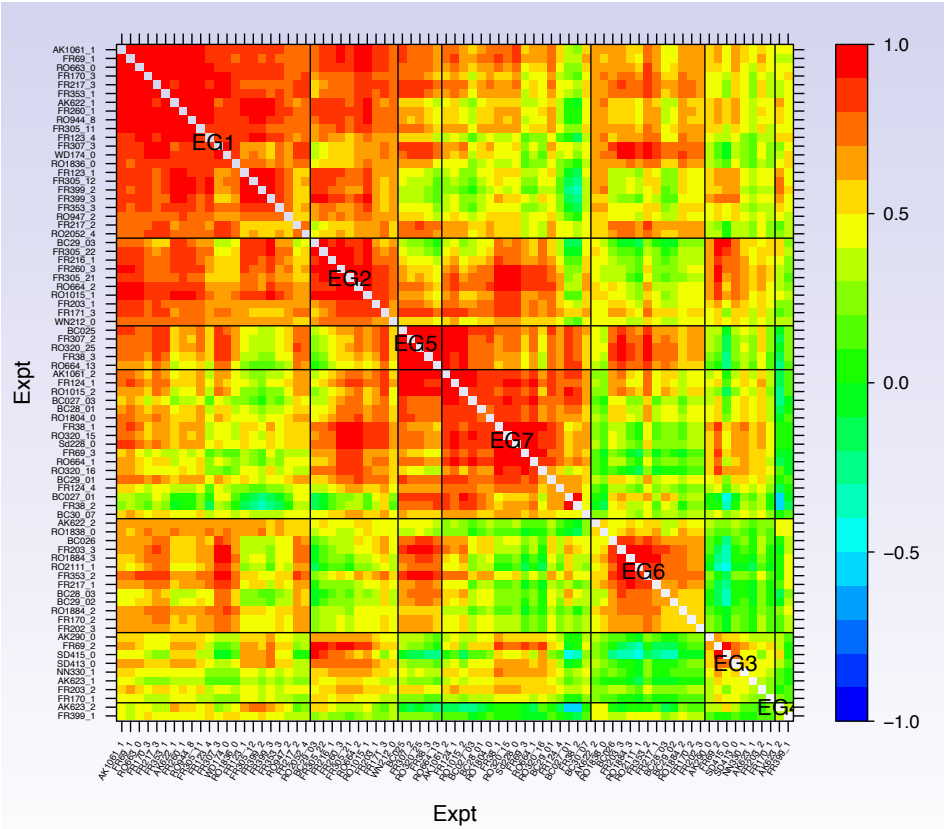
- Lack of/poor connectedness between trials is our largest problem in genetic evaluation.
- It means we **can't** compare some genotypes to each other.
- It also means that many genotypes have been tested under a small subset of environments
- One of the priorities of the RPBC breeding plan is to expand coverage and connectedness. We have ramped up trial installation for the last 5 years.

Additive genetic correlation matrix

Estimated



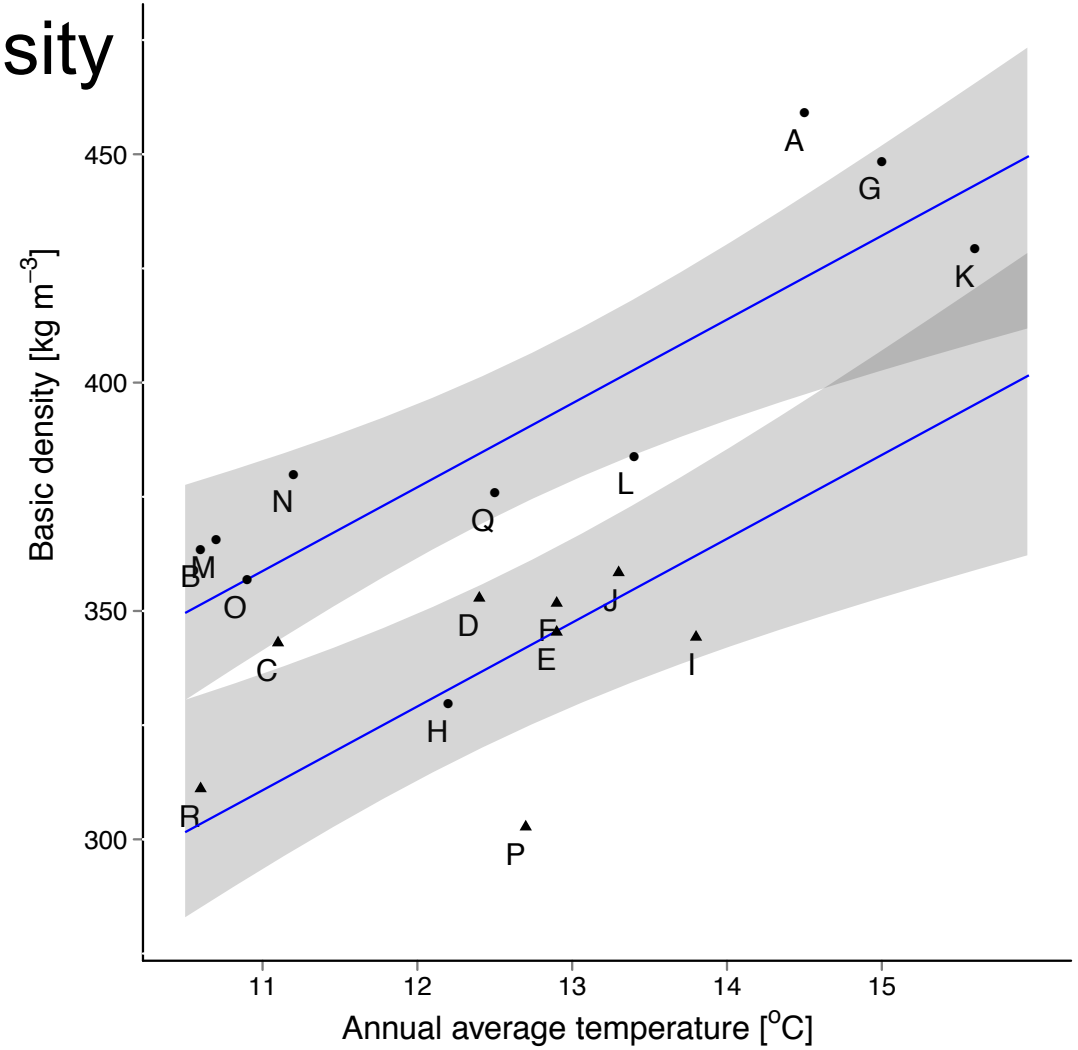
Sorted by clusters of correlation



Cullis, Smith, Jefferson & Thompson 2013. Implementation strategy for RPBC breeding values incorporating GxE

HOW DO WE EXPLAIN THESE CORRELATIONS?

Some times we have simple explanations for GxE; for example, scale effect of temperature on basic density



Rings 6-10

Rings 1-5

Site means based on 23,000+ increment cores in 18 sites

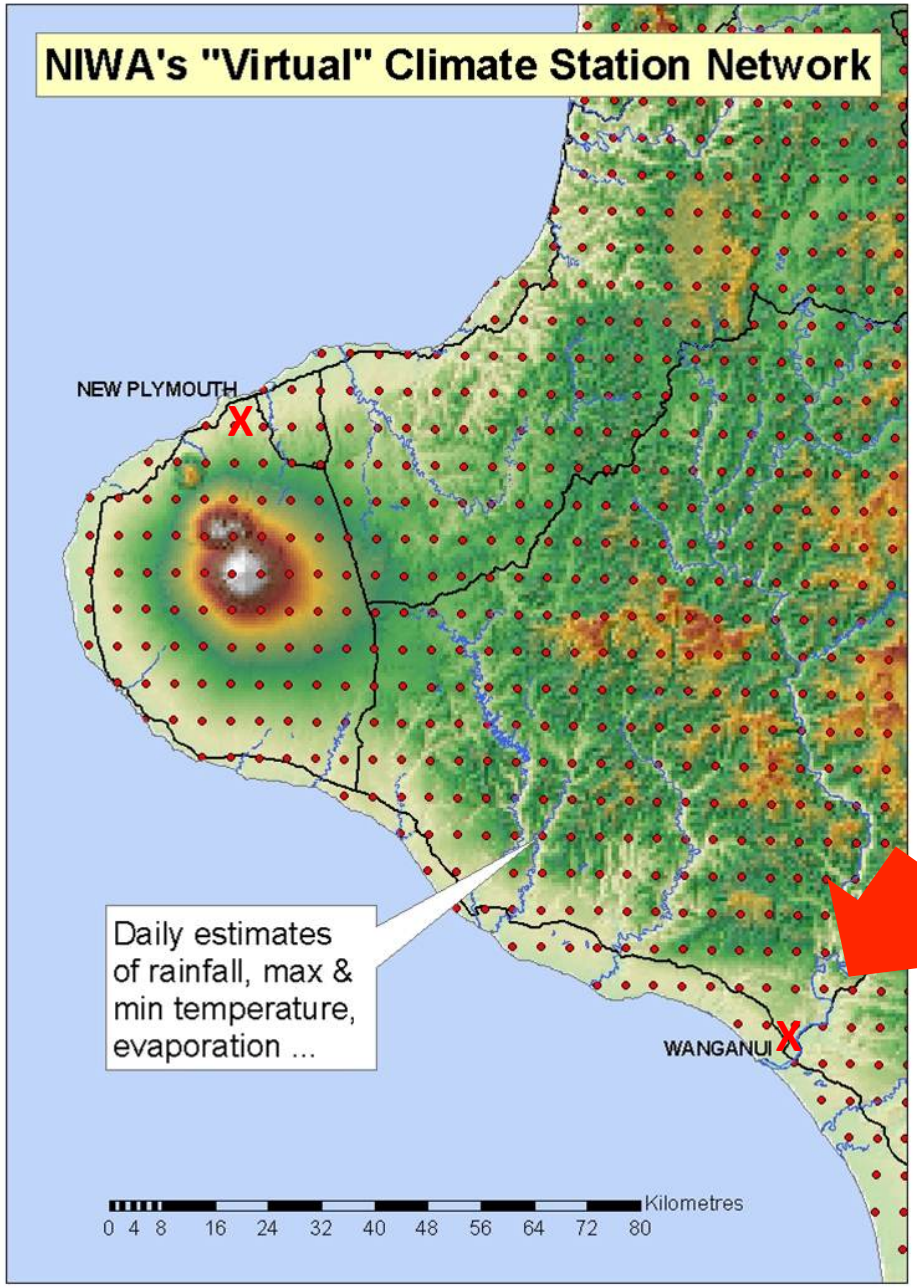
$R^2 = 0.75$

Apiolaza 2012 Basic density of radiata pine in New Zealand: genetic and environmental factors. TGG 8: 87-96

In contrast, for growth we don't know the drivers of GxE. Several attempts:

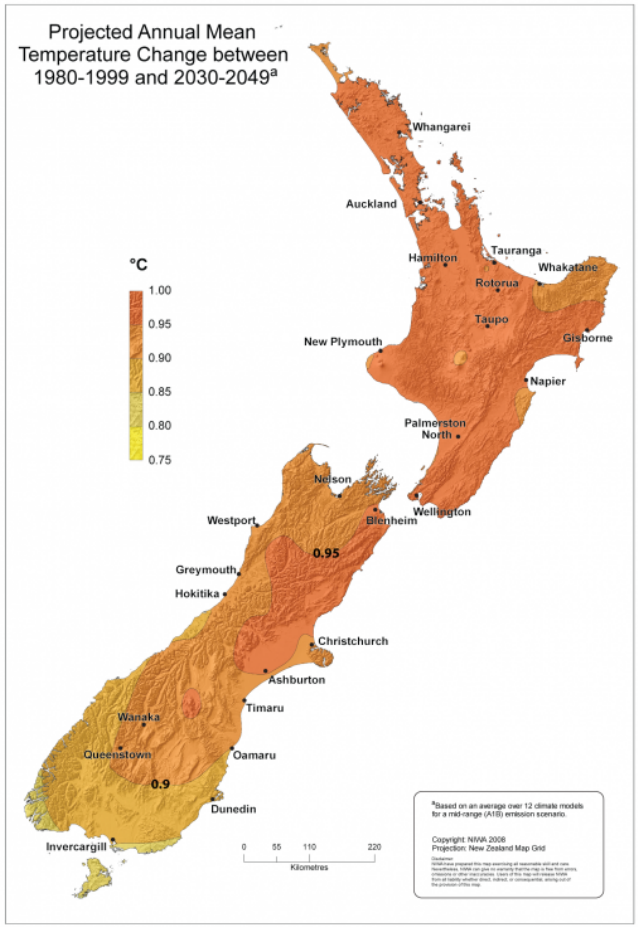
McDonald & Apiolaza 2008-9
Raymond 2010
Ivkovic et al 2012-14
Cullis & Jefferson 2013-14

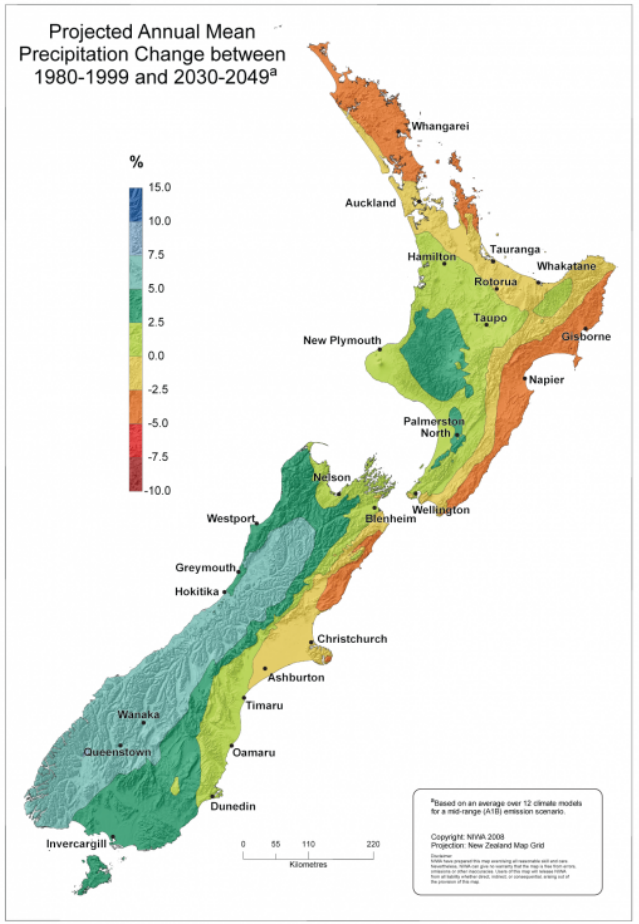
One **probable** cause: we have soil and climate data at the wrong scale (both in space and time)



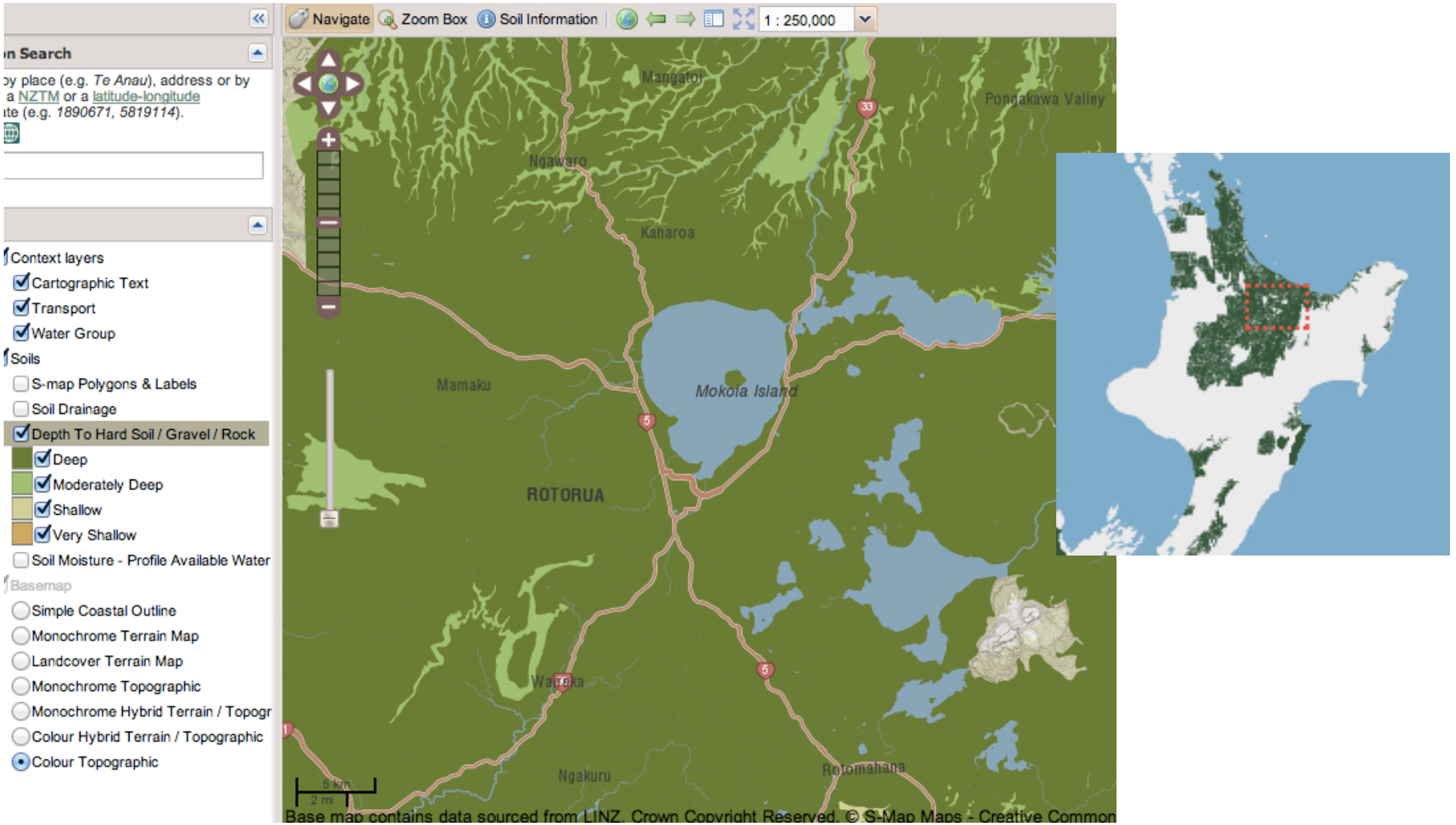
Example of 5 x 5 km grid
Red crosses indicate **actual** weather stations used in the interpolation

On top of that, climate will be different in our next rotation



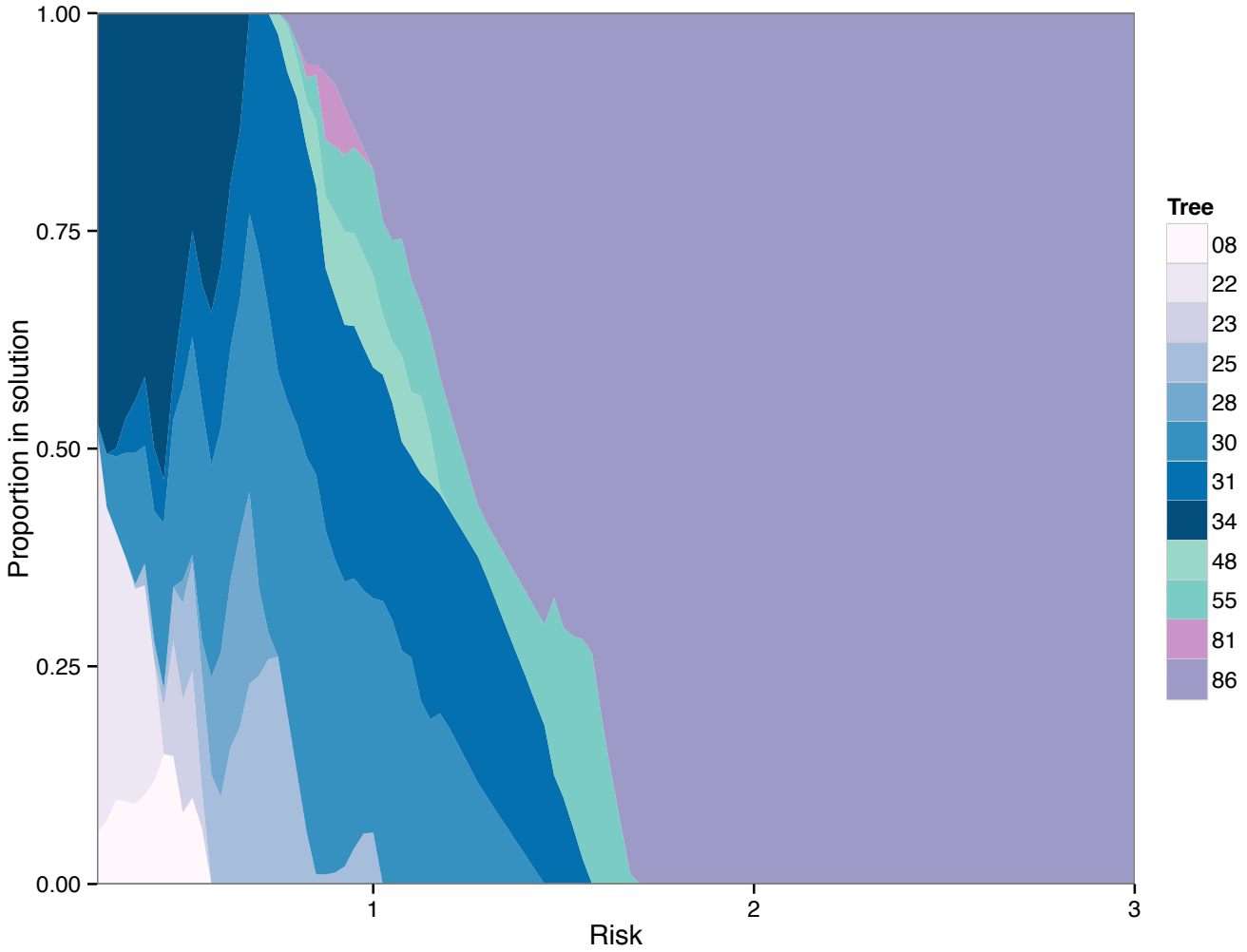


And, even worse, very poor resolution for soils



In summary, we have no first-hand, reliable information on environmental variables that drive GxE interaction for growth

An understanding of deployment environments permits adjusting our approach to risk



Apiolaza, L.A. and Alzamora, R.M. 2013. Building deployment portfolios for genotypes under performance instability. Silva Fennica 47(1): 901

Final remarks

- Today we can run a single-stage multivariate national evaluation (with Factor Analytic & Reduced Animal models) using ASReml-R.
- We are not yet able to explain the environmental factors driving GxE for growth.
- Probable cause: **wrong scale** environmental data.
- One of our priorities should be to invest in high resolution descriptions of the environment for our best, better-connected trials (**moving from G to E**).

Thanks to

- The New Zealand Radiata Pine Breeding Company and its members, who have supported my work since 2006, both with data and funding
- Paul Jefferson & Brian Cullis for early access to the results of the latest genetic evaluation.
- The conference organizers for kindly inviting me to present today.