Comparing Analyses:

Simon J. Greenhill (simon@simon.net.nz).

ARC Centre of Excellence for the Dynamics of Language, Australian National University.

Department of Linguistic and Cultural Evolution, Max Planck Institute for the Science of Human History.

We should now have two completed analyses – cpacific-ctmc-relaxed and cpacific-ctmc-strict. If you haven't gotten this far, then let me know and I'll give you a copy of my analyses.

`[ ] Make sure you have both analyses`

Open `Tracer` and load both log files from these two analyses into it.

`[ ] Load both log files into Tracer`

Comparing parameter estimates across analyses:

`Tracer` allows us to compare across different analyses quite easily. Let's look at the estimate age of these languages again:

1. Select both trace files in the `Trace Files` panel (top left)
2. Select the `TreeHeight` parameter in the `Traces` panel (on the left).
3. Look on the `Estimates` tab – you should see the vital statistics of both analyses for this parameter.

Summary Statistic	cpacific-ctmc-relaxed.log	cpacific-ctmc-strict.log	
mean	3261.7677	4266.1801	
stderr of mean	44.1795	2.4541	
stdev	1138.942	443.1098	
variance	1.2972E6	1.9635E5	
median	3123.1946	4233.3019	
mode	n/a	n/a	
geometric mean	3073.8223	4243.6193	
95% HPD Interval	[1291.3273, 5382.5411]	[3435.6589, 5142.3212]	
auto-correlation time (ACT)	1.3542E5	2760.7022	
effective sample size (ESS)	664.5961	32600.7641	

`[ ] Which analysis is closer to our expected age of ~3200?`

1. Select the `Marginal Prob. Distribution` tab.
2. Down the bottom of the screen turn on the `Legend`, and then set `Color By` to `Trace file`.

`[ ] What are the differences/similarities between the two estimated ages?`

Which analysis is best?

As a rule of thumb, the analysis with the smallest value for `posterior` (closer to zero) is better.
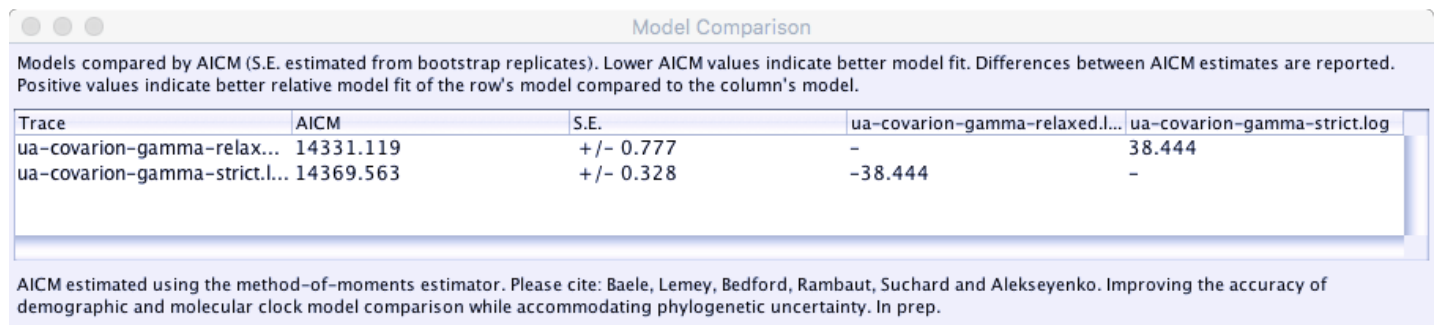
`[ ] Which analysis is better?`

But it's better if we use a formal model testing procedure to find which analysis is the best for these data. The easiest method at the moment is to calculate the `AICM` (Akaike Information Criteria through MCMC). The model with the *lowest* AICM is better, and differences in AICM scores > 7 provide `strong evidence` of one model over the other (Burnham & Anderson '98).

Note The AICM does not tell you that a model is *correct*, just that it is a better fit to the data.

To do this, in `Tracer` select `Analysis -> Model Comparison` from the file menu. Then make sure that `likelihood` and `aicm` are selected.

Calculating the AICM takes forever. If you were going to publish the results then you should make sure that `Bootstrap replicates` is set to 1000 so you get a good estimate. However, for our case we're going to set it to 5.

Here's the results from an analysis I ran earlier this week on `Uto-Aztecan` languages. Note that the first analysis (ua-covarion-gamma-relaxed) has a lower AICM than the other one by 38.444 units. This is substantial support for the first analysis being a better fit to the data.

Run the AICM on your data:

`[ ] Which analysis is better?`

`[ ] By how much?`

The AICM is the simplest way to compare models at the moment, but the current gold standard is an approach called `Path-Sampling` or `Stepping-Stone` (Beale et al. '02) but it needs more work and can take a lot of time to run. For publication you should run a Path-Sampling comparison following this tutorial http://beast2.org/2014/07/14/path-sampling-with-a-gui/.

Subgrouping of Polynesian Languages:

It turns out that there's a recent debate about how the Polynesian languages are sub-grouped. For a long time the accepted classification of these languages has followed Andy Pawley ('66, '67) with some revisions by Jeff Marck ('00).
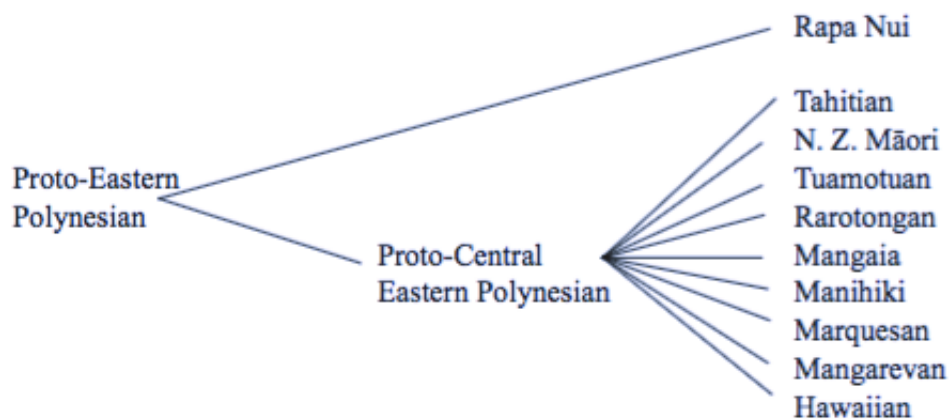
This is what the Pawley-Marck classification looks like:

FIGURE 1. EASTERN POLYNESIAN LANGUAGES
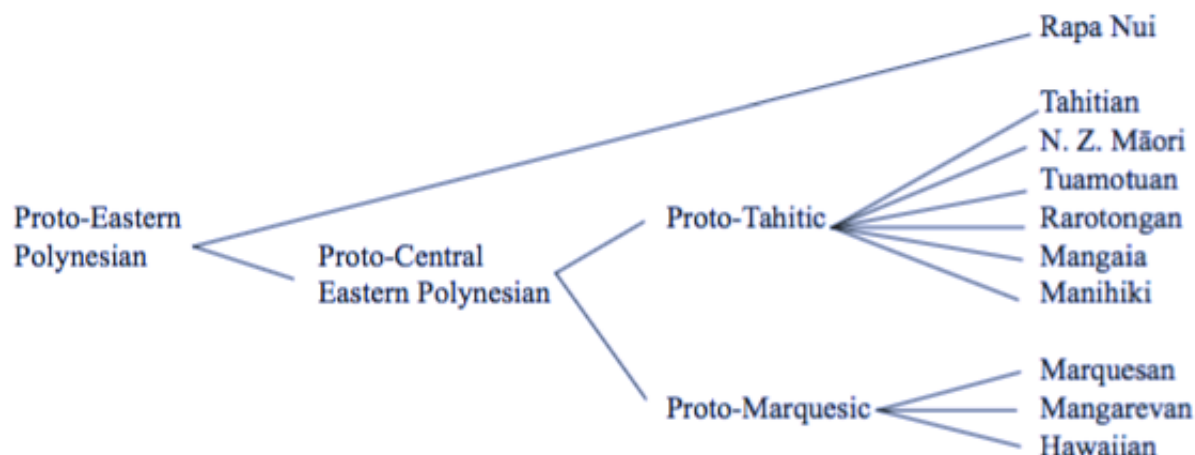
However, a more recent study by Mary Walworth ('14) has argued that the evidence for both `Proto-Tahitic` and `Proto-Marquesan` is weak and we'd be better off with this picture:

FIGURE 2. PROPOSED NEW TREE FOR EASTERN POLYNESIAN
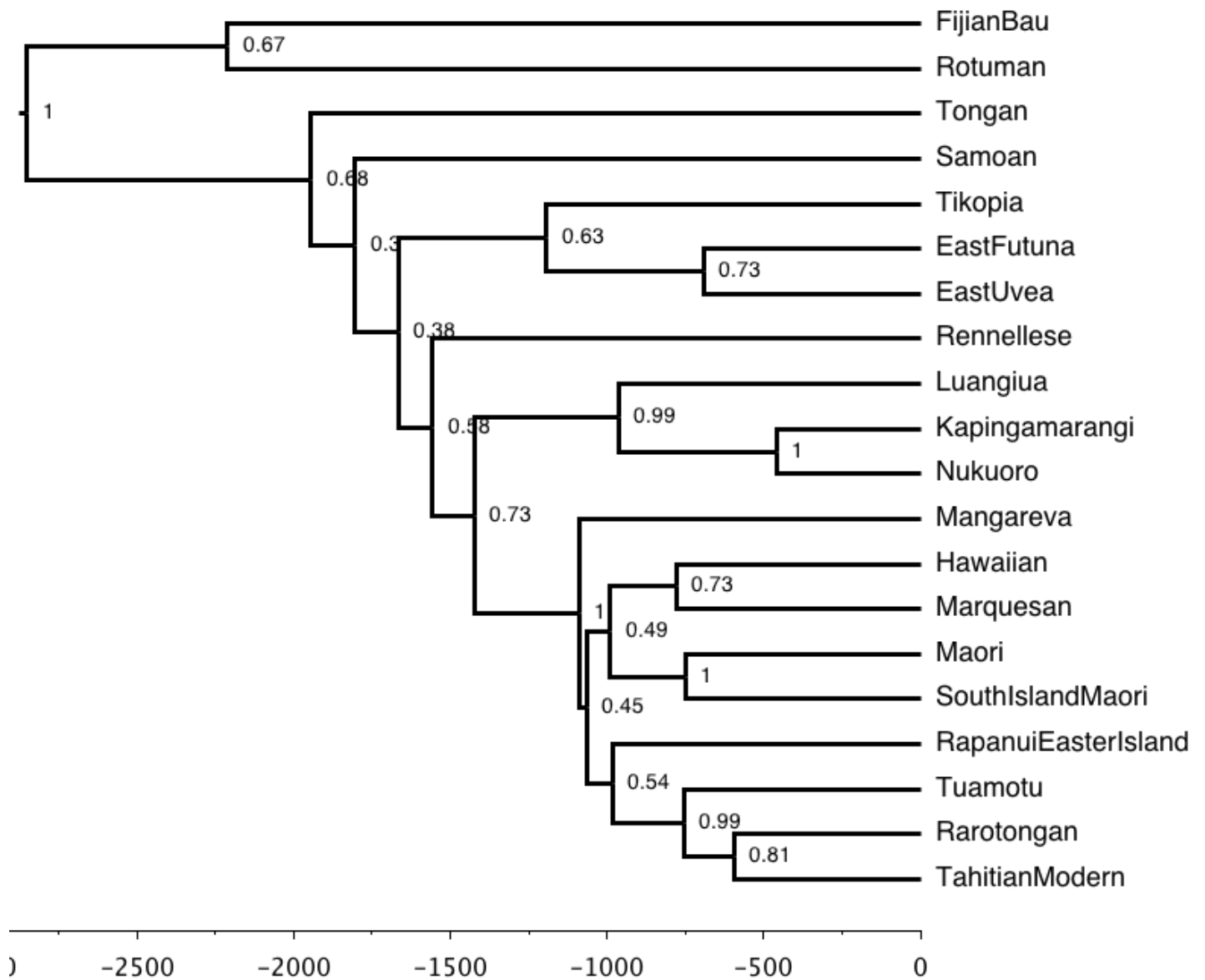
In our results we have the following Tahitic languages:

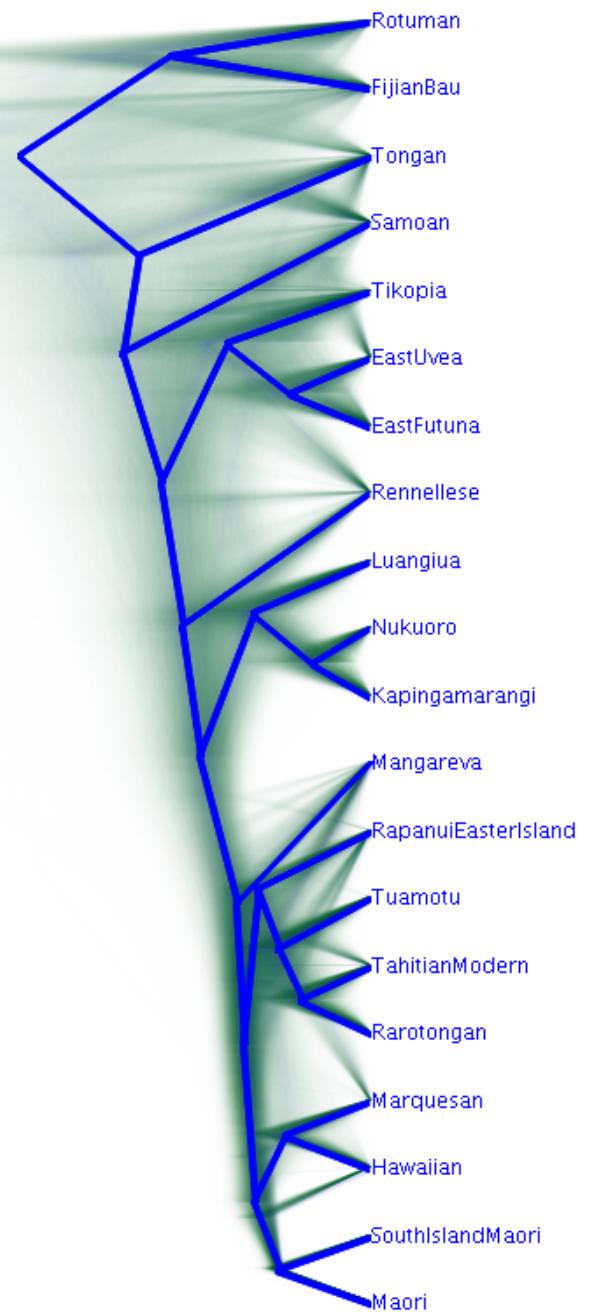- Tahitian, Maori, SouthIslandMaori, Tuamotuan, and Rarotongan.

… and the Marquesic languages:

- Marquesan, Mangarevan, and Hawaiian.

Our results from the best fitting model say this:

[ ] Which hypothesis do we support?

Conclusions:

I hope to have shown you how to:

- do some quality control checks on your data
- visualise data and interpret a Neighbor-Net
- set up a phylogenetic analysis to estimate some trees
- process this analysis to check for stability.
- make inferences about timing of Pacific settlement
- make inferences about the subgrouping of these languages.

Discussion:

1. here we've applied these methods to reasonably messy set of languages – lots of dialect chains and language contact. Do you think it worked? Was I too optimistic? When is it not appropriate to use these methods?

2. how could we make these methods better?

3. what about *your* data. Are there any questions that phylogenetic tools can help you answer?