

Pelos gráficos entendi que (não fiz nenhuma análise para saber se os dados são significativos):

1. O desvio padrão é baixo o que me leva a crer que existe pouca variação na população. Para ver se isso é verdade pensei em fazer um gráfico da variação entre os indivíduos a cada geração populacional.
 - Se realmente tiver pouca variação pensei em aumentar a taxa de mutação, ou utilizar um sistema auto-adaptativo (mas a baixa variação vem desde o início... Entendo que essa técnica é válida quando a variação vai se perdendo com o tempo) ou algum outro método.
2. Para Kanto e Kansai o desempenho da versão com listas é superior ao gamodel. Não sei se significativo.
3. Em Kanto o comportamento dos dois modelos, com profundidade < 60 km, é estranho (será que é algum erro?).
4. Em geral a versão com lista tem desempenho melhor quando temos menos terremotos. Ou seja em Kansai e Kanto é melhor que em Tohoku e EastJapan. Assim como o desempenho é melhor com profundidade 25km, depois com 60km e a pior é com 100km.
5. Já o gamodel sofre menos com essas mudanças.
6. No Kanto a situação é curiosa. O desempenho do ga é menor que a lista pra profundidade <25, é melhor quando profundidade < 60 e depois piora de novo.
7. Reparei que em geral praticamente não há variação da média do fitness por volta da geração 80.

Resumindo, entendo que a versão de lista para profundidades < 25km e profundidade < 60km tem desempenho melhor que o gamodel. Para profundidade < 100km perde. Portanto acho que vale a pena explorar o quão significativo são os dados.

O tempo gasto em geral do método é menor quando lidamos com menos terremotos. Hoje eu ficaria com terremotos de profundidade 25km para termos métodos mais rápidos.

Sinto que precisamos aumentar a diversidade.

Podemos diminuir a quantidade de avaliações dos métodos uma vez que a partir da geração 80 os dados pouco mudam.