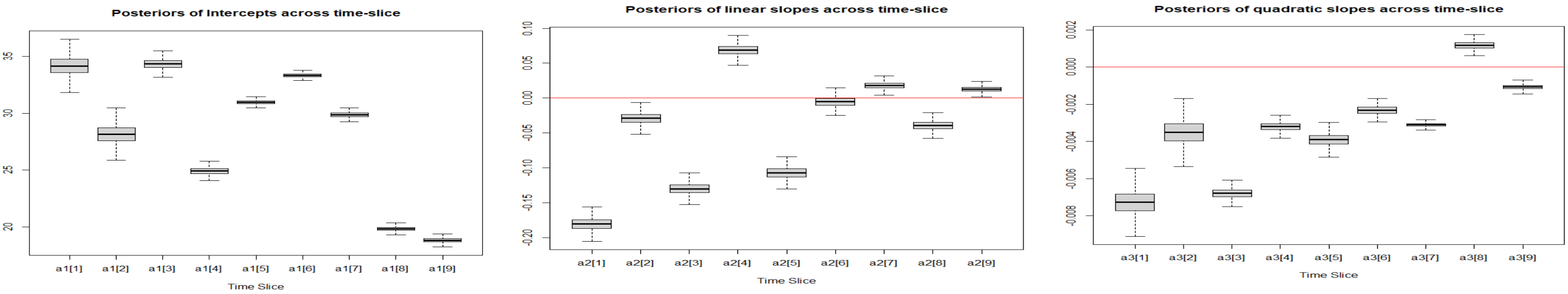
Let Y_{ij} denote the MAT estimate j at the time slice i and X_j the corresponding paleo latitude. We deploy a quadratic regression with random slopes and heteroskedastic effects since it is natural that different times slices will have different polynomial slopes. From the plot of the data, it is evident that the variance of the measurements varies with the paleo latitude X_j. We discern heteroskedasticity and spatial autocorrelation. MAT measurements on the same latitude form clusters , and the variance at different clusters changes, while it varies nonlinearly as we move away from latitudes in which MAT clusters. This is a common pattern across all the time slices. We consider the following models:

- **M1**: Varying slopes as fixed effects and homoscedastic errors
  $Y_{ij} \sim \mathrm{N}(a_{1i} + a_{2i}X_j + a_{3i}X_j^2, \sigma^2)$ with uninformative priors $a_{ki} \sim \mathrm{N}(0,1000)$, $\sigma^2 \sim Inv\Gamma(0,1,0.1)$
  for $i = 1{:}9,\ j = 1{:}7953,\ k = 1{:}3$
- **M2**: Random slopes with spatial heteroskedasticity
  $Y_{ij} \sim \mathrm{N}(a_{1i} + a_{2i}X_j + a_{3i}X_j^2, \sigma_j^2)$ with $\sigma_j^2 = \sigma_0^2 e^{\frac{(X_j - X_{j-1})}{c}}$ to capture that measurements on the same latitude will exhibit the same scale of variance. The random slopes are $a_{ki} \sim \mathrm{N}(b_k, \sigma_k^2)$ and the priors are: $c \sim \mathrm{U}[0, \max(\mathrm{X}) - \min(\mathrm{X})]$, $b_k \sim \mathrm{N}(0,1000)$, $\sigma_0^2 \sim Inv\Gamma(0,1,0.1)$, $\sigma_k^2 \sim Inv\Gamma(0,1,0.1)$
- **M3**: Random slopes with polynomial heteroskedasticity
  $Y_{ij} \sim \mathrm{N}(a_{1i} + a_{2i}X_j + a_{3i}X_j^2, \sigma_j^2)$ with $ln(\sigma_j^2) = \mu + a_{4i}X_j^2 + a_{5i}X_j^3$, random slopes $a_{ki} \sim \mathrm{N}(b_k, \sigma_k^2)$ for $k = 1{:}5$,and uninformative priors $\mu \sim \mathrm{N}(0,1000)$, $b_k \sim \mathrm{N}(0,1000)$, $\sigma_k^2 \sim Inv\Gamma(0,1,0.1)$.

We use 3-fold cross-validation (2,651 quantitative obs /fold), DIC and WAIC as selection methods to compare fits:

Model	BIAS	MSE	MAD	COVERAGE	DIC	WAIC
M1	-0.0026351	23.43644	3.608211	0.0126996	47,593	-47,445
M2	0.0138869	23.49312	3.597822	0.0144600	47,510	-47,357
M3	-0.0273762	24.62638	3.587526	0.0153401	46,808	-46,575

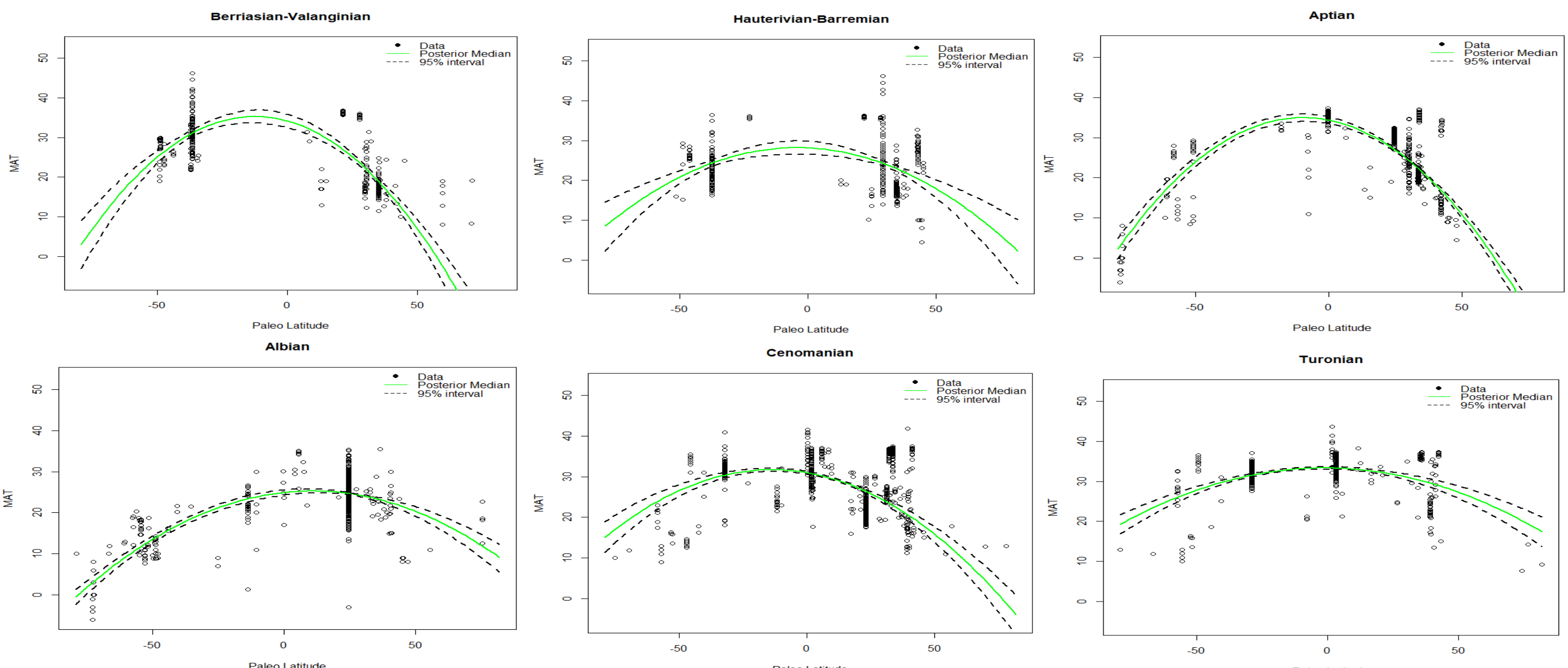
We prioritize MAD over MSE in judging model adequacy due to the scale of the data. M3 has the highest absolute bias. DIC prefers M3, while WAIC prefers M1. M3 has the smallest MAD and the highest coverage, though coverage is poor for all models. Based on MAD, Coverage and DIC, M3 is the optimal one to carry the rest of the analysis.
The posteriors of the regression coefficients for M3 are:

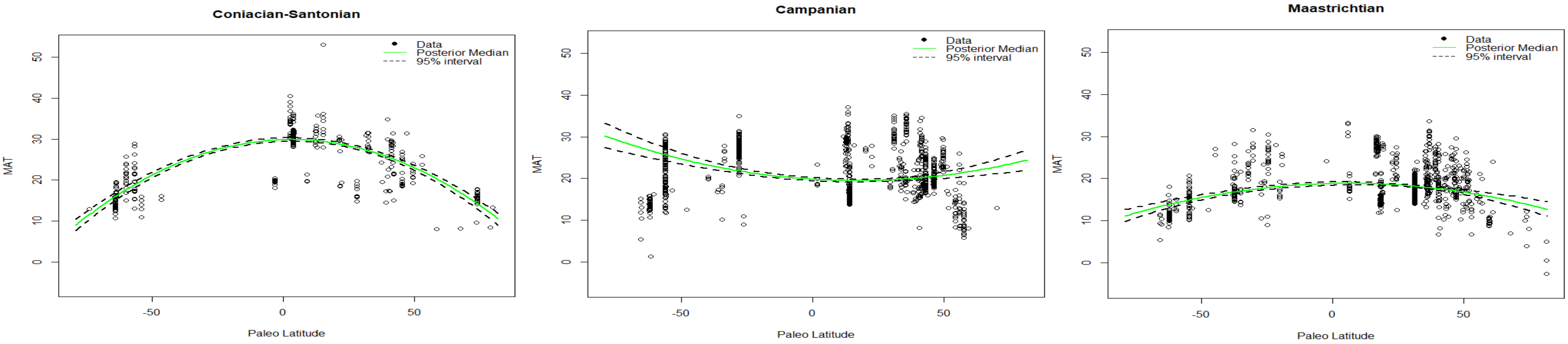


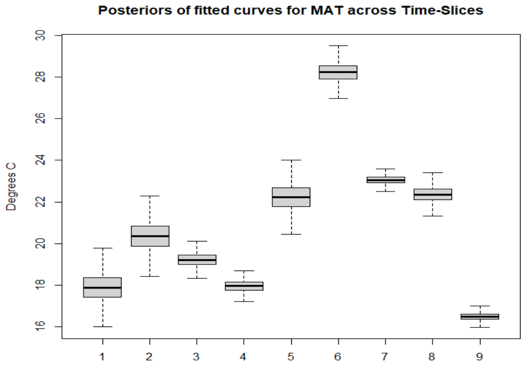
All intercepts, linear and quadratic regression coefficients are statistically significant except from the linear coefficient a_{26} of time slice 6 whose posterior includes 0. The convergence of the MCMC chains looks great for all the coefficients. The mean effective sample size is 10,236. The trace plots, ACF plots and ESS are included in the appendix.
We perform goodness-of-fit tests by computing the following Bayesian p-values:

Model M3	Mean Y	SD Y	Min Y	Max Y	Range Y
Bayesian p-value	0.43792	0.99996	0.00004	0.92592	1.00000

Apart from the mean, all other p-values are approximately 0 or 1. Thus, the model does not fit well. To enhance performance, we can include higher order polynomials and model the autocorrelation through an autocovariance matrix with spatial covariance and (i,j) element: $\sigma^2 e^{\frac{(X_i - X_j)}{c}}$. **1E)** Estimated curves using the posterior median:



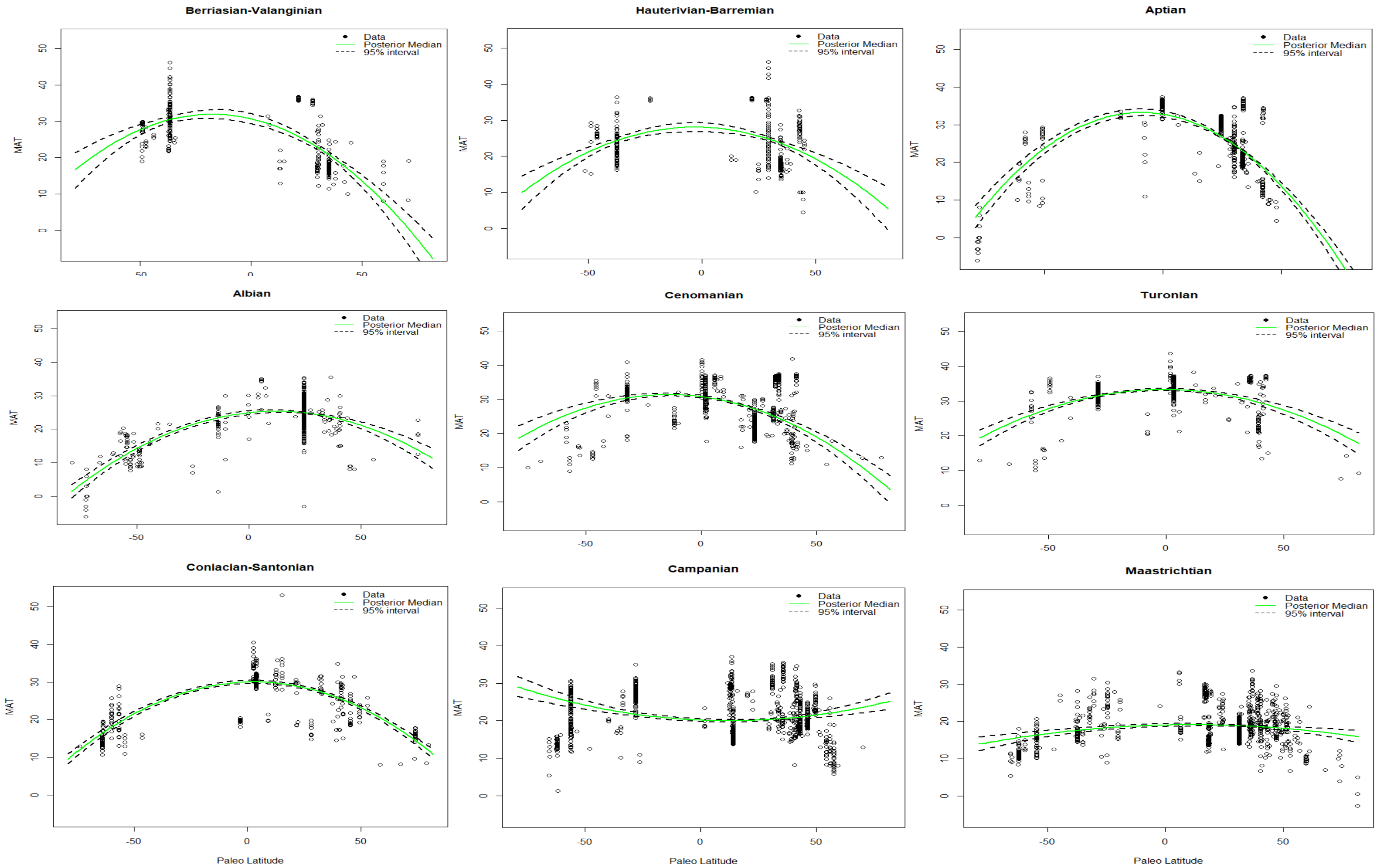




According to the plot of the posteriors of the fitted curves to the left, there is strong evidence that MAT changes significantly across time-slices since the estimated posteriors change levels.

For the mixed data analysis we use the same model (M3) as before. To accommodate interval censoring we use the ***dinterval*** distribution of JAGS for the censored observations as: I[j] $\sim dinterval(Y_{n+j}, [min(Y_{n+j}), max(Y_{n+j})])$ where j=1:J (J=2120) is a censored measurement, and n is the number of observed measurements. This produces a likelihood $P(I|Y) = \begin{cases} 1, if\ min(Y_{n+j}) \leq Y_{n+j} \leq max(Y_{n+j}) \\ 0,\ o.w \end{cases}$. If I=1, the censored Y_{n+j} is sampled from its PPD with the restriction of belonging to the given interval.

Estimated curves using the posterior median for the mixed data fit:



The posterior predictive checks for the model estimated using the total of 10,073 mixed observations yield:

Model M3	Mean Y	SD Y	Min Y	Max Y	Range Y
Bayesian p-value	0.99956	0.86872	0.00288	0.85180	0.99980

We observe a slight improvement in the p-values of SD, Min and Max, but the p-value of the Mean worsens compared to the quantitative data analysis. Further, the linear regression coefficient a_{26} still comes out insignificant and the estimated curves do not change much. For the given model, there is no evidence that the inclusion of the interval data improves the analysis.