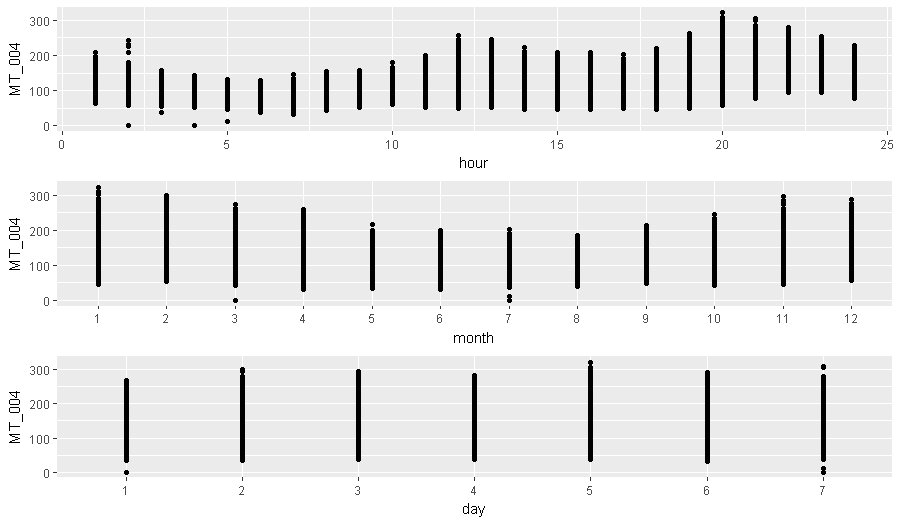
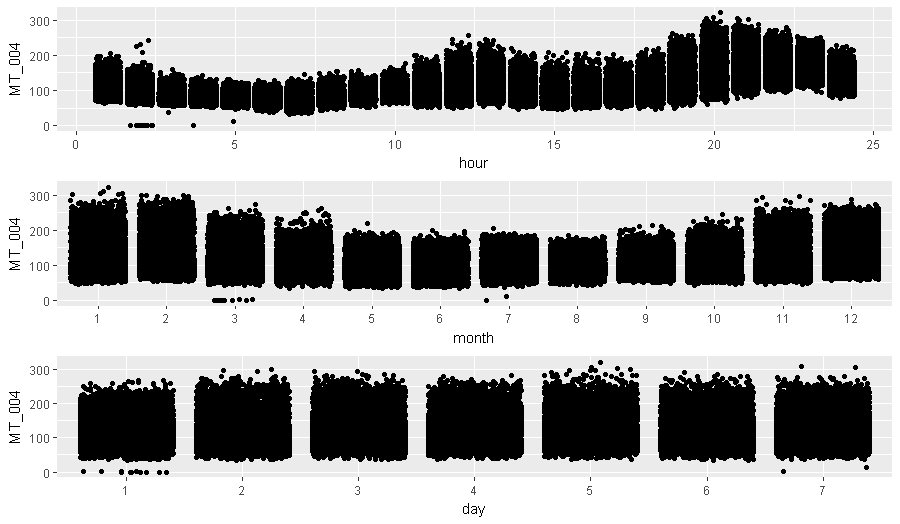
GAM BASELINE DIAGNOSTICS

*Core Model (Main Effects)*

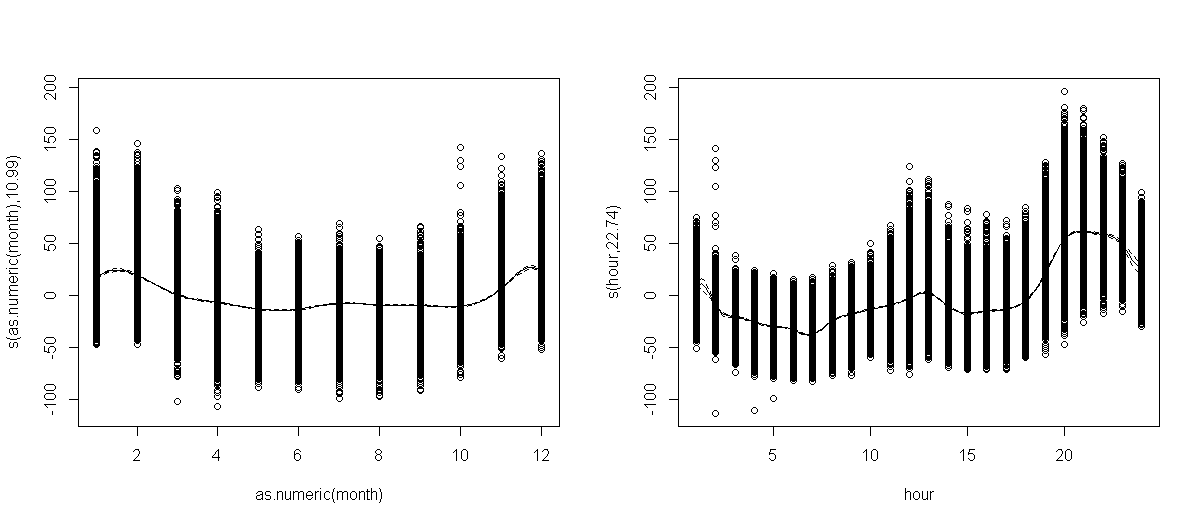
NOTE: This model does required a significant amount of data because many coefficients must be estimated. I would suggest at least two complete months. Until then a simpler model like loess may be used for baseline calculation.

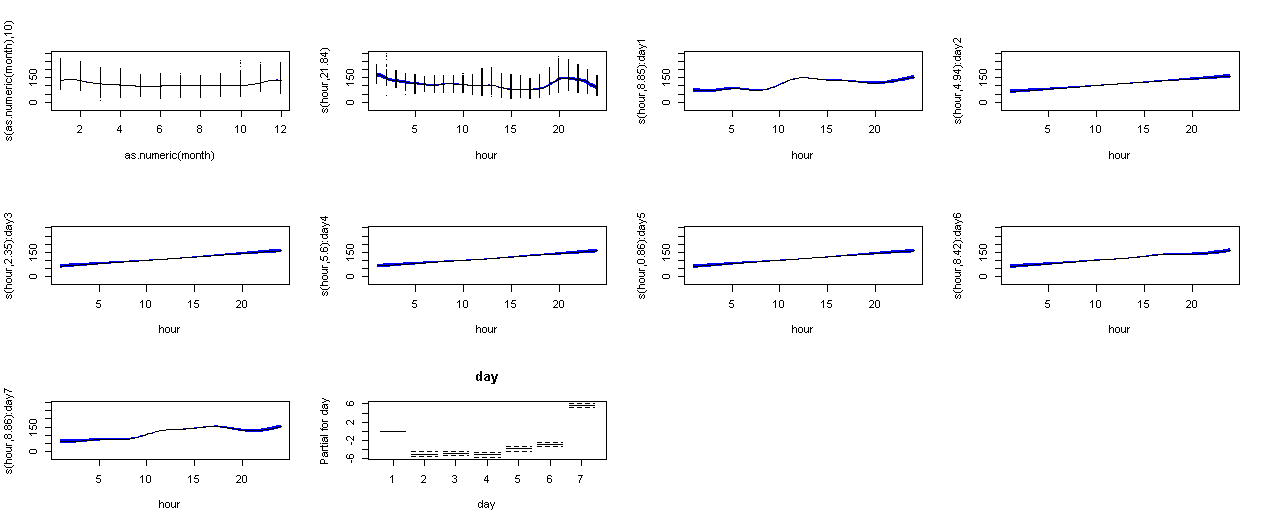


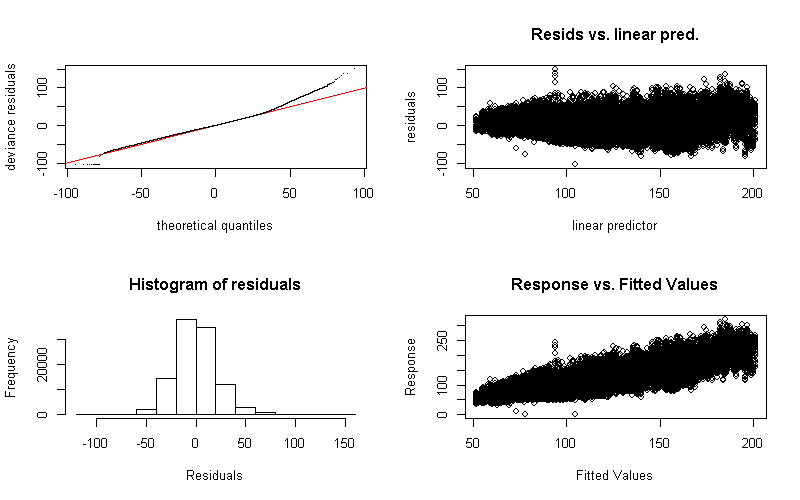


## BASE MODEL: PARTIAL EFFECT PLOTS & DIAGNOSTIC PLOTS

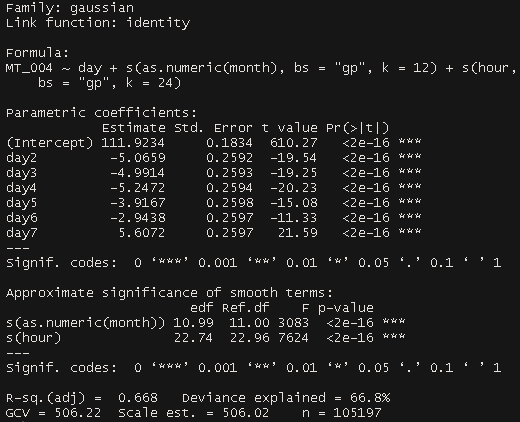
MT\_004 <- gam(MT\_004 ~ day + s(as.numeric(month), bs = "gp", k = 11) + s(hour, bs = "gp", k = 23), data = structured\_data)





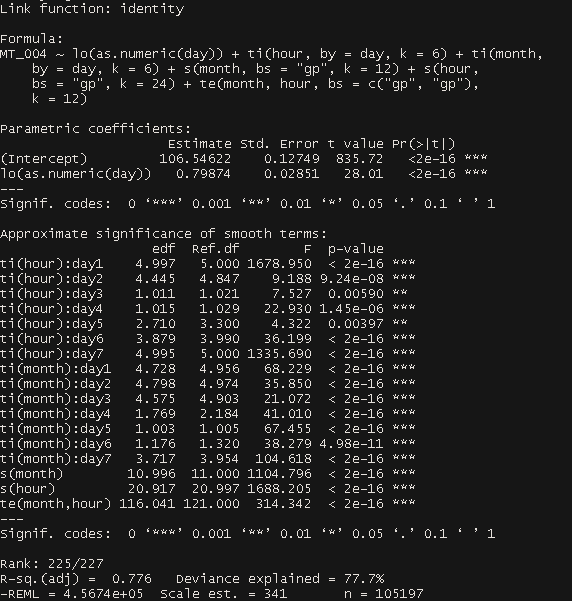


Inputting day and month as factors gives use the same error rate as only day and applying spline to month. It does appear that the spline is more appropriate given the partial effect plots at top.



# *Proposed Model Including all Identified Interactions*

MT\_004 <- bam(MT\_004 ~ lo(as.numeric(day)) + ti(hour, by = day, k = 6) + ti(month, by = day, k = 6) + s(month, bs = "gp", k = 12) + s(hour, bs = "gp", k = 24) + te(month, hour, bs = c("gp","gp"), k = 12), data = structured\_data, method = "REML")



NOTE: Although, coding the main effect of day as categorical/factor (ie dummy code for each level) produces a higher deviance explained it appears to give poorer out of sample predictions than using the lowess method on variable day.

## PROPOSED MODEL: PARTIAL EFFECT PLOTS & DIAGNOSTIC PLOTS

