Published in final edited form as:

Stat Med. 2018 June 15; 37(13): 2108–2119. doi:10.1002/sim.7627.

A Three Level Bayesian Mixed Effects Location Scale Model with An Application to Ecological Momentary Assessment Data

Xiaolei Lin, Robin J. Mermelstein, and Donald Hedeker*

John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Abstract

Ecological Momentary Assessment (EMA) studies usually produce intensively measured longitudinal data with large numbers of observations per unit, and research interest is often centered around understanding the changes in variation of people's thoughts, emotions and behaviors. Hedeker et al.¹ developed a two level mixed effects location scale model that allows observed covariates as well as unobserved variables to influence both the mean and the within-subjects variance, for a two level data structure where observations are nested within subjects. In some EMA studies, subjects are measured at multiple waves and within each wave subjects are measured over time, Li and Hedeker² extended the original two level model to a three level data structure where observations are nested within days and days are then nested within subjects, by including a random location and scale intercept at the intermediate wave level. However, the three level random intercept model assumes constant response change rate for both the mean and variance. To account for changes in variance across waves, as well as clustering attributable to waves, we propose a more comprehensive location scale model that allows subject heterogeneity at baseline as well as across different waves, for a three level data structure where observations are nested within waves and waves are then further nested within subjects. The model parameters are estimated using Markov Chain Monte Carlo methods. We provide details on the Bayesian estimation approach and demonstrate how the Stan statistical software can be used to sample from the desired distributions and achieve consistent estimates. The proposed model is validated via a series of simulation studies. Data from an adolescent smoking study are analyzed to demonstrate this approach. The analyses clearly favor the proposed model and show significant subject heterogeneity at baseline as well as change over time, for both mood mean and variance. The proposed three level location scale model can be widely applied to areas of research where the interest lies in the consistency in addition to the mean level of the responses.

Keywords

Ecological Momentary Assessment; mixed effects; variance modeling; Bayesian sampling

*Correspondence to: Donald Hedeker, Department of Public Health Sciences, The University of Chicago. 5841 South Maryland Ave MC2000, Chicago, IL 60637-1447. DHedeker@health.bsd.uchicago.edu.

1. Introduction

Modern data collection procedures, such as ecological momentary assessments (EMA), allow researchers to study outcomes with high volatility by repeated sampling of subjects' behaviors and experiences in real time and subjects' natural environments³. Typically these procedures involve self-reported data collection from individuals over the course of hours, days, and weeks, thus yield relatively large numbers of observation per subject⁴. A particular interest in EMA studies is to identify factors that affect the within subject variance of the intensively measured outcomes, in addition to the overall mean levels⁵. Due to the hierarchical nature of EMA data, random subject effects are usually included in statistical models to account for the correlation among repeated measures for a given subject⁶. Hedeker et al.¹ developed a mixed effect location scale model that includes an additional random subject effect in the error variance, thus allowing subject variation in terms of both the mean and variance of the intensively measured outcomes. Random effects in both the mean and variance model can be useful in distinguishing the residual variation from unobserved subject-level variables, thus providing more accurate standard errors and valid statistical inference⁷.

EMA studies are sometimes conducted at multiple measurement waves, resulting in a three level data structure: observations nested within waves, and waves in turn nested within subjects⁸. For example, a person's mood can be assessed multiple times at each wave and the subject can be followed up at multiple waves. There are three possible sources of mood variation for this type of data: variation between subjects, variation within subject but between waves, and variation within subject within wave. Ignoring any possible sources of variation would lead to misspecification of the correlation structure and invalid statistical inference. Li and Hedeker² proposed a three level mixed effect location scale model that includes random subject and day intercepts for both the mean and within variance of the outcome. Kapur et al.⁹ proposed a similar Bayesian mixed effect location scale model for multivariate outcomes at one EMA wave. However, these models assume that subjects change with a constant rate in terms of both mean and variance. This assumption can be easily violated, especially in psychological and behavioral studies, as subjects almost always exhibit heterogeneous trajectories across time¹⁰. Using the above mood example, subjects are likely to have different mood variability at baseline, and over time, some can become more consistent while others become more erratic. Rast et al.¹¹, Leckie¹², and Goldstein et al.¹³ all presented a two level mixed effect location scale model that includes random intercept and slope for both the location and scale model, allowing for heterogeneous trajectories across time. Therefore, a three level model that treats observations within waves within subjects while accounting for subject heterogeneity at baseline and over time for both mean and variance will provide a more comprehensive utilization of the data as well as address more specific questions of interest. However, estimation of such general models with relatively large numbers of random effects using likelihood-based methods can be prohibitive due to computational and numerical complexity¹⁴.

In this article, we propose a Bayesian mixed effect location scale model for three level data structures, where observations are nested within waves, and waves further nested within subjects. The proposed model extends the conventional three level mixed effect regression

model by including random subject intercept and slope as well as random wave intercept for both the mean and within-subject variance of the outcome. At the mean level, the proposed model allows subjects to have heterogeneity in their baseline responses as well as different growth rates over time. Similarly at the variance level, subjects are allowed to exhibit different variation at baseline and the variation can also change differentially over time. Both the subject and wave level heterogeneity can be explained by observed covariates as well as unobserved variables through specification of random effects. Furthermore, the random location and scale effects are allowed to be correlated. The proposed model is estimated using a Bayesian approach. Specifically, Markov Chain Monte Carlo sampling methods are used to generate samples from the joint posterior distribution, and parameter estimates and credible intervals are obtained by summarizing the corresponding distributions¹⁵. We will demonstrate how Stan (an open source Hamiltonian Monte Carlo sampler) and the Hamiltonian Monte Carlo algorithm can be used to achieve consistent parameter estimates and we provide a detailed syntax example in the Supplemental Materials¹⁶. The model is validated via a sequence of simulation studies against several reduced models. Finally the proposed three level model is applied to an EMA adolescent smoking study, where the interest is on identifying risk factors associated with high mood variation as well as exploring the possible mood trajectories.

2. Motivating Adolescent Smoking Study Example

The data that motivate the development of the Bayesian three level mixed effect location scale model are from an EMA adolescent smoking study. In the study, 461 adolescents from 9th and 10th grade were recruited. The average age of the participants is 15.6, with the minimum being 14.4 and maximum 16.7. They carried hand held devices for 7 days at each measurement wave, during which they responded to random interviews (approximately 5 times per day) or event recorded any episodes of smoking. At each prompt or smoking episode, participants were asked to answer questions, including location, activities, companionship, mood and other psychological measurements. The study was conducted at 6 waves: baseline, 6 months, 15 months, 2 years, 5 years, and 6 years. Data were collected on age, gender, beep type (smoking event vs random prompt), and positive affect (measure of positive mood). Because of the interest in comparing responses across waves from random prompts vs smoking events, subjects were included if they were measured on at least two waves and had at least two smoking events at each wave, resulting in a sample size of 254 subjects.

Among all the subjects, 51.6% were female, and on average subjects were followed up at three waves with 36 to 51 prompts (including smoking episodes) per wave during the entire study span. A total of 24490 random prompts and 8087 smoking events were obtained, with an approximate average of 96 random prompts and 32 smoking episodes per subject. For the analyses reported, a three level structure of observations (level1) within waves (level 2) within subjects (level 3) was considered.

The outcome is the measure of subjects' positive affect (PA), which consists of the average of several mood items rated from 1 to 10: I felt happy, I felt relaxed, I felt cheerful, I felt confident, and I felt accepted by others. Thus, higher PA levels indicate better mood. The

interest is to see whether subjects tend to have higher and more consistent positive affect after smoking compared to random prompts. We are also interested in differentiating the between subject and within subject between wave effect from the within subject within wave effect. That is, the effect of smoking when comparing different subjects (between subject), the same subject at different waves (within subject between wave), and the same subject at the same wave but different occasions (within subject within wave). Since the three variables contain different information in characterizing subjects' smoking behavior, it is useful to include the decomposed variables in the model and investigate their relative statistical and clinical significance so that further interventions can be done at that level⁸. Investigation into the PA showed that subjects exhibit different trends across waves in terms of both mean and variability, as shown in Figure 1 and Figure 2.

3. Method

Suppose there are $k = 1, \dots, n_j$ observations nested within $j = 1, \dots, n_i$ waves, and waves are then nested within $i = 1, \dots, n$ subjects. Let y_{ijk} denote the outcome for subject i measured at wave j and occasion k . The conventional three level mixed effect model can be expressed as

$$Y_{ijk} = X_{ijk}^T \beta + Z_{ijk}^T \gamma_i + V_{ijk}^T \nu_{ij} + \epsilon_{ijk} \quad (1)$$

where X_{ijk} is the $p \times 1$ vector of regressors (typically including a column of "1" for the intercept) which can contain either subject, wave or occasion level variables, and β is the corresponding vector of regression coefficients. Z_{ijk} (usually a subset of X_{ijk}) is the vector of regressors for random effect γ_i , and γ_i is the vector of random subject effect, indicating the influence of individual i on his/her repeated mood assessments. Similarly, V_{ijk} (again, usually a subset of X_{ijk}) is the vector of regressors for random effect ν_{ij} , where ν_{ij} represents the vector of random wave effect, indicating the influence of wave j on subject i 's repeated mood assessments.

For the EMA adolescent smoking study example, the outcome Y_{ijk} is the positive affect for subject i at wave j and occasion k . Since we are interested in differentiating the within subject within wave effect from the between subject as well as the within subject between wave effects, we will decompose the occasion level variable smk_{ijk} (1 for smoking event and 0 for random prompt) into subject, wave and occasional level variables.

$$\overline{smk}_i = \sum_j \sum_k smk_{ijk} / \sum_j K_{ij}, \quad \overline{smk}_{ij} = \sum_k smk_{ijk} / K_{ij} - \overline{smk}_i, \quad \widetilde{smk}_{ijk} = smk_{ijk} - \overline{smk}_{ij} \quad (2)$$

Here, K_{ij} is the number of observations for subject i at wave j ; $\overline{smk}_i$ is the decomposed subject level variable and represents the average (proportion) of smoking events for subject i ; $\overline{smk}_{ij}$ is the decomposed wave level variable and represents the deviation of average smoking events at wave j relative to the subject level average $\overline{smk}_i$; $\widetilde{smk}_{ijk}$, which is computed

as the deviation of smoking events at occasion k relative to the subject's wave level average, represents the pure occasion level smoking effect adjusted for his/her subject and wave level average. All three variables will be included in the mean model. For random subject effects, both a random intercept and a random slope over wave will be included since there is interest about subject heterogeneity both at baseline and trajectories over time. So Z_{ijk} will be two dimensional and consists of a column of 1 and wave indicator $wave_{ij}$. Correspondingly, $\gamma_i = \{\gamma_{0,i}, \gamma_{1,i}\}$, with $\gamma_{0,i}$ being the random subject intercept indicating the influence of subject i on his/her baseline mood, and $\gamma_{1,i}$ being random subject slope indicating the influence of subject i on how fast or slow his/her mood changes over time. Since our data have a three level structure with an intermediate wave clustering, an additional random wave effect should be included. For wave, only a random intercept will be considered to indicate the possible influence of wave on subjects' repeated mood assessments: even for the same subject, the mood can be different at different waves and the difference cannot be fully explained by the observed wave level variables. As a result, V_{ijk} will be a column of 1 and v_{ij} is of dimension 1. Therefore, the mean model for the adolescent mood study example can be expressed explicitly as

$$Y_{ijk} = \beta_0 + \beta_1 male_i + \beta_2 \overline{smk}_i + \beta_3 \overline{smk}_{ij} + \beta_4 \widetilde{smk}_{ijk} + \beta_5 wave_{ij} + \gamma_{0,i} + \gamma_{1,i} wave_{ij} + \nu_{ij} + \varepsilon_{ijk}$$

(3)

The random effects γ and ν are referred to as location random effects since they influence the mean or location of the outcome. Both γ and ν are assumed to be normally distributed with constant variance covariance structure Σ_γ and σ_ν^2 , and independent of each other. The size of the diagonal elements in Σ_γ indicates the amount of between subject variability, while size of σ_ν^2 indicates the amount of the within subject between wave variability. The random error ε_{ijk} are usually assumed to be normally distributed with constant variance σ_ε^2 . However, since σ_ε^2 represents the amount of variability that exists within subjects and within waves, by assuming σ_ε^2 constant, we are assuming that the within variance does not vary for different subjects or waves. This assumption can be easily violated in practice, especially for psychological and behavioral studies, where subjects almost always exhibit variation in terms of the consistency in their responses. One approach to relax this assumption is to additionally model σ_ε^2 by another mixed effect model through a log-linear representation

$$\log(\sigma_{ijk}^2) = \alpha_0 + \alpha_1 male_i + \alpha_2 \overline{smk}_i + \alpha_3 \overline{smk}_{ij} + \alpha_4 \widetilde{smk}_{ijk} + \alpha_5 wave_{ij} + \lambda_{0,i} + \lambda_{1,i} wave_{ij} + \tau_{ij} \quad (4)$$

Similar to the mean model 2, the within variance model contains both fixed effects α and random effects $\{\lambda, \tau\}$. In addition to the observed variables $\{male_i, \overline{smk}_i, \overline{smk}_{ij}, \widetilde{smk}_{ijk}, wave_{ij}\}$ that can influence the variability of the outcome for certain subject at certain waves, there can also be unmeasured variables contributing to how consistent/erratic the outcome measurements could possibly be. Ignoring the unobserved information would lead to invalid inference about the variance parameters. This motivates the inclusion of random scale effects in the within variance model 4. At subject level, $\lambda_{0,i}$ is the random scale intercept and indicates the influence of subject i on his/her mood variability at baseline, and $\lambda_{1,i}$ the random scale slope, indicates the influence of subjects on how the variability changes over time. For example, some subjects may start off with relatively consistent responses (small within variance at baseline), but over time their responses become more and more erratic (positive slope on the within variance over wave), while others may follow some different patterns. This heterogeneity among subjects in terms of the variance trajectories can be captured by the random subject scale intercept λ_0 and slope λ_1 . At the wave level, only a random scale intercept τ will be considered to account for the possible effect of wave on the within variance.

An intuitive visualization of the model mechanics is shown in Figure 3. There are two hypothetical subjects: subject 1 has both increasing positive affect and mood variation, while subject 2 has increasing positive affect but diminishing mood variation across wave. We can also visualize the wave effect as different waves exhibit different mean as well as variation. The different patterns suggest different mood trajectories as well as disease prognostics from a psychological perspective, which our proposed model is able to capture.

There are six random effects, consisting both subject and wave levels, in terms of both the mean and within variance of the outcome. The population distribution of these random effects is similar to the ordinary mixed effects models in that random subject effects can be possibly correlated but should be independent of the random wave effects. In addition, the random location effects are allowed to be possibly correlated with the random scale effects, as extreme mean values are often accompanied with more consistent variance due to ceiling or floor measurement effects. The distributional assumption for the six random effects can be expressed as

$$\begin{bmatrix} \gamma_{0,i} \\ \gamma_{1,i} \\ \lambda_{0,i} \\ \lambda_{1,i} \end{bmatrix} \sim \mathcal{N}_4 \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\gamma_0}^2 & \text{cov}_{\gamma_0, \gamma_1} & \text{cov}_{\gamma_0, \lambda_0} & \text{cov}_{\gamma_0, \lambda_1} \\ \text{cov}_{\gamma_0, \gamma_1} & \sigma_{\gamma_1}^2 & \text{cov}_{\gamma_1, \lambda_0} & \text{cov}_{\gamma_1, \lambda_1} \\ \text{cov}_{\gamma_0, \lambda_0} & \text{cov}_{\gamma_1, \lambda_0} & \sigma_{\lambda_0}^2 & \text{cov}_{\lambda_0, \lambda_1} \\ \text{cov}_{\gamma_0, \lambda_1} & \text{cov}_{\gamma_1, \lambda_1} & \text{cov}_{\lambda_0, \lambda_1} & \sigma_{\lambda_1}^2 \end{bmatrix} \right) \quad (5)$$

$$\begin{bmatrix} \nu_{0,ij} \\ \tau_{0,ij} \end{bmatrix} \sim \mathcal{N}_2 \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\nu_0}^2 & \text{cov}_{\nu_0, \tau_0} \\ \text{cov}_{\nu_0, \tau_0} & \sigma_{\tau_0}^2 \end{bmatrix} \right) \quad (6)$$

4. Model estimation

To estimate the model parameters, Bayesian approaches are favored against maximum likelihood methods which usually involve heavy numerical integration and approximation of the first and second order partial derivatives¹⁷. For a typical Newton Raphson algorithm to achieve MLEs, one would need to integrate the conditional likelihood over the joint distribution of all random effects in order to compute the marginal likelihood. As a result, the computational load and complexity increases exponentially with the number of random effects, making the estimating procedure infeasible for models with relatively large numbers of random effects¹⁸. Bayesian approaches, on the other hand, perform the estimation by drawing Markov Chain Monte Carlo samples from the joint posterior distribution given the prior that reflects our belief about the parameters before collecting the data¹⁹. Various sampling algorithms can be used, including a mixture of Gibbs sampling, Metropolis-Hastings and Hamiltonian Monte Carlo²⁰. Parameter estimates and credible intervals can then be obtained by taking the point estimates and corresponding intervals associated with the posterior, thus avoiding the computational issues associated with numerical integration²¹. Given flat priors and enough MCMC samples, the Bayesian approach will yield consistent parameter estimates²².

We've devised a MCMC sampling algorithm where Metropolis-Hastings algorithms is nested within Gibbs sampling. However, a better approach can be taken using the Stan statistical software, since it can better deal with the trade off between step size and acceptance rate by reducing the correlation between successive samples using a Hamiltonian evolution and target values with a higher acceptance rate than the observed probability distribution²³. Both the Metropolis-Hastings-Gibbs sampling algorithm and Stan implementation details are provided in the supplemental materials. The Hamiltonian Monte Carlo sampling uses improper uniform priors (uniform on $(-\infty, +\infty)$) for regression coefficients β and α , improper bounded uniform priors (uniform on $(0, +\infty)$) for random effect variances σ_{subj}^2 and σ_{wave}^2 , and LKJ (Lewandowski, Kurowicka and Joe) priors for random effect correlation matrix.

5. Simulation Study

To validate the proposed model as well as the estimation procedure, a simulation study was conducted. A series of 100 data sets, each with 10000 observations (100 subjects, each subject measured at 10 waves and 10 occasions within each wave) were generated under the three level location scale model with three covariates ($X_i \sim \mathcal{N}(0, 1)$, $X_{ij} \sim \mathcal{N}(0, 1)$, $X_{ijk} \sim \mathcal{N}(0, 1)$). The true parameter values for the six random effects variances/covariances are

$$COV \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \lambda_0 \\ \lambda_1 \end{pmatrix} = \begin{bmatrix} 1.0 & -0.1 & 0 & 0 \\ -0.1 & 0.25 & 0 & 0 \\ 0 & 0 & 0.25 & -0.025 \\ 0 & 0 & -0.025 & 0.065 \end{bmatrix}, \text{ and } COV \begin{pmatrix} \nu_0 \\ \tau_0 \end{pmatrix} = \begin{bmatrix} 1.0 & 0.25 \\ 0.25 & 1.0 \end{bmatrix}. \text{ For each generated}$$

data set, a series of four candidate models were considered: a two level (subject and occasion level) mixed effect regression model with heterogeneous variance (MRM HV), a two level mixed effect location scale model (MLS), a three level (subject, wave and occasion level) mixed effect regression model with heterogeneous model, and the proposed three level mixed effect location scale model. The first three models are considered to be reduced models relative to the last one since they ignore either the clustering due to the intermediate wave or unobserved variables in the variance.

Two level MRM HV:

$$Y_{ik} = X_{ik}^T \beta + \gamma_{0,i} + \gamma_{1,i} wave_{ik} + \varepsilon_{ik} \quad (7)$$

$$\sigma_{\varepsilon,ik}^2 = \exp(W_{ik}^T \alpha) \quad (8)$$

Two level MLS:

$$Y_{ik} = X_{ik}^T \beta + \gamma_{0,i} + \gamma_{1,i} wave_{ik} + \varepsilon_{ik} \quad (9)$$

$$\sigma_{\varepsilon,ik}^2 = \exp(W_{ik}^T \alpha + \lambda_{0,i} + \lambda_{1,i} wave_{ik}) \quad (10)$$

Three level MRM HV:

$$Y_{ijk} = X_{ijk}^T \beta + \gamma_{0,i} + \gamma_{1,i} wave_{ij} + \nu_{0,ij} + \varepsilon_{ijk} \quad (11)$$

$$\sigma_{\varepsilon,ijk}^2 = \exp(W_{ijk}^T \alpha) \quad (12)$$

Three level MLS:

$$Y_{ijk} = X_{ijk}^T \beta + \gamma_{0,i} + \gamma_{1,i} wave_{ij} + \nu_{0,ij} + \varepsilon_{ijk} \quad (13)$$

$$\sigma_{\varepsilon,ijk}^2 = \exp(W_{ijk}^T \alpha + \lambda_{0,i} + \lambda_{1,i} \text{wave}_{ij} + \tau_{0,ij}) \quad (14)$$

The four candidate models were compared in terms of both mean and variance parameter estimates as well as credible intervals. Bias, average 95% credible interval width (AIW) and average coverage rate (COV) out of 100 data sets were obtained to evaluate model performance. The results are represented in Table 1 and Table 2.

In Table 1, $\beta^{intercept}$, β^{subj} , β^{wave} , β^{obs} are the mean model regression coefficients for the intercept, subject, wave and occasion level covariates, respectively. The four models all did relatively well in estimating β as can be seen from the small bias. But they do perform different in terms of estimating the uncertainties associated with the coefficients: the three level models (three level MRM HV and three level MLS) produced wider and more correct intervals (1.6873/1.6993 and 1.8105/1.8039) for the intercept and wave covariate compared to the two level models (0.4050/0.3973 and 0.2103/0.2075). This is due to the fact that neither two level MRM HV nor two level MLS account for the possible unobserved variables at baseline or the intermediate wave level by including random intercept or random wave effect(s), which in turn over states the certainty around the point estimates. Although the point estimates in all four models show small bias, only the three level models yield credible intervals closer to the correct 95% level.

In Table 2, $\alpha^{intercept}$, α^{subj} and α^{wave} are the corresponding regression coefficients associated with the intercept, subject and wave level covariates in the log-linear representation of the error variance model. All three reduced models have $\alpha^{intercept}$ estimates biased upwards with narrower credible intervals and insufficient coverage. One explanation is that, when one omits the wave level covariates or random scale effects (or both) in the log linear error variance model, all the variations unexplained by the existing covariates has to be absorbed by $\alpha^{intercept}$, which makes $\alpha^{intercept}$ biased towards the population averaged effect rather than subject specific effects. Leckie¹² had similar findings regarding the variance model intercept in a two level random intercept location scale model. In terms of α^{subj} and α^{wave} , since neither MRMs included random scale effects, they produced narrower and incorrect credible intervals. Also, the two level MLS under covers α^{wave} due to the fact that it failed to include a random scale effect at the wave level.

In summary, none of the reduced models are comparable to the three level mixed effects location scale model in terms of unbiasedness and correct coverage. If one were to analyze a three level structure data sets where both location and scale random effects are present, using the reduced models would yield invalid statistical inference and arrive at possibly false positive results.

6. Application to adolescent smoking study

The proposed Bayesian three level mixed effect location scale model was applied to the EMA adolescent smoking study introduced in section 2. For comparison purposes, results from a three level mixed effect regression model, as well as a three level mixed effect

regression model with heterogeneous variance were also listed. The focus was on identifying risk factors associated with lowered and unstable mood assessments, with an special interest in separating the within subject within wave effect from the between subject and within subject between wave effects of smoking events vs random prompts. The outcome is positive affect, which is a measure of a subjects' positive mood as described in the motivating example section. The occasion level covariate *smk* (1 if the response is from a smoking event, or 0 if from a random prompt) was decomposed into subject, wave and occasion level variables as described earlier since we are interested in identifying the most significant level of smoking effect. Wave is a continuous variable with values from 0 (baseline) to 6 (6 years after baseline); to facilitate computation, we made one unit equal to five calendar years so that it takes values from 0 to 1.2. In the supplemental materials, we have included R code to simulate a similar three level data set as well as run the Stan program from R.

Results are summarized in Table 3, for both mean and variance models. Since parameters were estimated using a Bayesian approach, Hamiltonian Monte Carlo samples were obtained from the posterior distributions for all parameters. The point estimates were obtained as the mean of the posterior distribution for regression coefficients β and α (since their posterior distributions are approximately symmetric), and as the mode of the posterior for random effect variances σ^2 (since their posterior distributions are skewed and mode would be most similar to the MLEs if one were to do likelihood estimation methods). The 95% credible intervals were bounded by 2.5% and 97.5% quantiles of the posterior for all parameters. The first two columns list the parameter estimates and corresponding credible intervals of the three level mixed effect regression model (MRM) which assumes homogeneous error variance and includes random subject location intercept and slope as well as random wave location intercept; the 3rd and 4th columns list results of the three level MRM which has the same random effects specification, but allows the error variance to depend on observed covariates; the final two columns list results of the proposed three level mixed effect location scale model (MLS), which, in addition to the random location effects, also includes the random scale effects and further allows the error variance to depend on both observed and unobserved covariates. The top two panels list regression coefficients for the mean β and within error variance α , with α on the natural log scale; the third panel lists the variances and covariances of the random effects, both for location and scale; and the bottom lists the model selection criteria, the expected log pointwise predictive density, or *elpd*, for all three models. *elpd_{LOO}* is a measure of how well the model fits the data controlling for the model complexity, and is often used for Bayesian model comparison²⁴. According to Vehtari et al.²⁴, *elpd_{LOO}* is preferred over deviance information criterion (DIC) since it evaluates the likelihood over the entire posterior distribution, works for singular models and is invariant to parametrization. Higher *elpd_{LOO}* indicates better model fit adjusting for the model complexity.

From Table 3, all random effect variances in the three level MRM, three level MRM with heterogeneous variance, and three level MLS are estimated to be greater than 0. But since the variance parameters are bounded, a preferred way to judge the significance would be to compare the *elpd_{LOO}* of the current model to those without corresponding random effects. The model selection criteria *elpd_{LOO}* strongly favors the three level MLS relative to either the three level MRM or three level MRM with heterogeneous variance. This provides clear

evidence that the homogeneous error variance assumption is violated, and observed information is insufficient to explain the amount of variation either at the subject level or at the wave level. Subjects do exhibit heterogeneity in terms of both mood and mood variation, and the heterogeneity in mood variation can be explained by some unmeasurable variables that are absorbed into random subject and wave effects. Specifically, subjects mood variation differs significantly at baseline and changes with different rates over time. The negative covariance between the scale intercept and slope indicates that subjects with more erratic mood at baseline exhibit greater mood stabilization across time, though this is not quite statistically significant as the credible interval includes zero.

When comparing the mean effects β among the models, all three models give similar results except for smk^{obs} , where the two MRMs yield a larger marginal effect compared to MLS. For all three models, smk^{obs} and $wave$ are seen to be statistically significant. For smk^{obs} , the point estimate is positive with 95% credible interval not including 0. This suggests that if we compare the same subject at the same wave, the subject tends to have better mood after a smoking event compared to after a random prompt. For $wave$, the point estimate and credible interval are both positive, indicating that across waves, subjects' mood tends to improve. Although the 95% credible interval contains 0, the results for smk^{subj} and smk^{wave} suggest that, for different subjects, heavier smokers tend to have higher mean mood; for the same subject across different waves, his mood tend to be better after a smoking event compared to a random prompt. Similar results among the three models suggests that, if the main interest is in the mean effects or changes in the mean, the ordinary MRM, MRM with heterogeneous error variance, and MLS all provide valid results.

When comparing the variance effects α among the models, the three level MRM assumes homogeneous error variance and thus only provides an intercept estimate; the three level MRM with heterogeneous variance, on the other hand, has a log linear representation of the error variance, and thus provides a point estimate and corresponding credible interval for each observed covariate. Additionally, the three level MLS further permits unobserved variables to affect the error variance by including the random scale intercept and slope, thus the three level MLS provides α for covariates as well as variances of the random scale effects. Results and conclusions from the latter two models differ, as can be expected based on the simulation study. Since the three level MRM with heterogeneous variance does not include the random scale effects, the point estimates for α might be reasonable, but the credible intervals will likely be too narrow. As can be seen from Table 3, the effects of $male$, smk^{subj} and smk^{wave} all tend to be significant by the three level MRM with heterogeneous variance, but not the three level MLS, due to the narrow credible intervals of the former. These positive effects are likely to be false positives since the three level MRM with heterogeneous variance tends to underestimate the uncertainty associated with α . Based on the three level MLS, smk^{obs} and $wave$ have negative effects on the variance and are seen to be statistically significant. For smk^{obs} , if we compare the same subject at the same wave, the subject tends to have more consistent mood after a smoking event compared to a random prompt. For $wave$, subjects' mood tends to become more consistent across time, as can be depicted in Figure 1. Although the 95% credible interval contains 0, the results for smk^{subj} and smk^{wave} suggest that, for different subjects, heavier smokers tend to have more stable mood; for the same subject across different waves, his mood tends to be more stable after a

smoking event compared to a random prompt. The dramatic differences in the results of α among the three models indicate that, neither MRMs provides adequate information or valid statistical inference if the main interest is centered around the variance effects or change of variation, in which case one should consider the proposed location scale model.

Data from four representative subjects were plotted to illustrate the subject and wave heterogeneity. In Figure 1, subject 1 was measured at all six waves while subject 2 was measured only at the last two waves. Subject 1 entered the study with relatively bad and unstable mood, but over time, his/her mood became better and more consistent. At the last wave, he/she provided very consistent high positive affect responses. This is consistent with the random subject scale effects estimates $(-0.41, -3.64)$, where the slope effect is estimated to be far below the population average. Subject 2, with the random scale slope estimated to be -0.25 , showed a somewhat similar pattern: he/she entered the study at wave 5 with unstable mood assessments, but became more consistent at wave 6. In Figure 2, subject 3 was measured at baseline and the last three waves, while subject 4 was measured at baseline and the last two waves. Subject 3 entered the study with relatively stable mood and then remained to be consistent throughout the study. Alternatively, subject 4, who entered the study with erratic mood assessments, then remained erratic until the end, which can also be depicted from his/her random scale intercept estimate $(1.36, 1.52)$, which is above the population average. These four subjects showed distinct patterns in terms of baseline mood variation as well as mood variation trajectories over time.

7. Discussion

In this article we have extended the existing two level mixed effects location scale model proposed by Hedeker et al.¹ to a three level structure and additionally allowed for multiple random location and scale effects. The three level mixed effect location scale model allows covariates to influence both the mean and within variance of the outcome, and thus relaxes the homogeneous error variance assumption. This model also includes random effects in both the mean and variance model, allowing variation in the outcome that cannot be fully explained by the covariates. The multiple random effects at the subject and wave levels allows variation in outcome trajectories among subjects and across waves, and provides more realistic assumptions as opposed to simpler random effect models. The magnitude of the random effect variance can help to reveal the degree to which heterogeneity is due to subjects and/or waves. Markov Chain Monte Carlo sampling methods were used to estimate the model parameters and to avoid numerical computation problems caused by the large number of random effects. Our example using the adolescent smoking data showed that subjects experience systematic mood variation at baseline as well as change over time.

The proposed model can be generalized to various research settings where the interest is in both the mean and variation of the outcome and where multiple levels of data clustering are present, such as smoking cessation²⁵ or substance addiction²⁶ studies. Since EMA studies produce relatively large number of observations per subject, the location scale model with both mean and variance modeling not only relaxes the constant error variance assumption but also permits more valuable information in terms of the outcome and subject (wave) heterogeneity. Furthermore, the proposed method can also be modified and used in a non-

EMA setting where data are collected from a series of hierarchical units instantly. For example, in clinical settings, glucose levels are often measured multiple times per day for type II diabetes patients at possibly multiple waves²⁷. The research interest often involves comparing the possible trajectories as glucose levels evolve with or without insulin pumps and thus infer the effectiveness of insulin pump therapies.

In this article, we only considered the possible effects of covariates on the within variance. However, one can also expand our model to additionally allow covariates to influence the between subject, as well as the within subject between wave variance²⁸. To do this, we need to include another set of between variance models. Specifically, let γ_i denote the random subject location effects, and Σ_{γ_i} be the variance covariance matrix of γ_i . Then the model for the diagonal elements in Σ_{γ} can be expressed as $\exp(X_i \times \eta)$, where X_i is the set of subject level covariates that have an effect on the between subject variance, and η is the corresponding regression coefficient. Similarly, we can include $\exp(X_{ij} \times \rho)$ to model the diagonal elements in the within subject between wave covariance matrix. Leckie¹² discussed the option for modeling the two by two between subject covariance matrix by specifying a log link function for the variances and inverse tanh link for the correlation. But it is trickier to extend well to higher order covariance matrices.

Our current work focuses on continuous outcomes only. Future work could therefore extend the current model and estimation framework to ordinal outcomes as well as count outcomes, by including a scale model representation for the overdispersion²⁹. Since ordinal and/or count outcomes generally provide less information compared to continuous outcomes, one might need to collect more data points in order to achieve relatively equal statistical power.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The project described was supported by National Heart Lung and Blood Institute (grant number R01 HL121330) and National Cancer Institute (grant number 5P01CA09862). The content is solely the responsibility of the authors and does not necessarily represent the official views of National Heart Lung and Blood Institute or National Cancer Institute. The authors thank Timothy Au for assisting in Stan programming.

References

1. Hedeker D, Mermelstein RJ, Demirtas H. An application of a mixed-effects location scale model for analysis of ecological momentary assessment (ema) data. *Biometrics*. 2008; 64:627–634. [PubMed: 17970819]
2. Li X, Hedeker D. A three-level mixed-effects location scale model with an application to ecological momentary assessment (ema) data. *Statistics in Medicine*. 2012; 31:3192–3210. [PubMed: 22865663]
3. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annual Review of Clinical Psychology*. 2008; 4:1–32.
4. Ebner-Priemer UW, Trull TJ. Ecological momentary assessment of mood disorders and mood dysregulation. *Psychological Assessment*. 2009; 21(4):463–475. [PubMed: 19947781]
5. Stone AA, Shiffman S. Ecological momentary assessment (ema) in behavioral medicine. *Annals of Behavioral Medicine*. 1994; 16(3):199–202.

6. Schwartz JE, Stone AA. Strategies for analyzing ecological momentary assessment data. *Health Psychology*. 1998; 17(1):6–16. [PubMed: 9459065]
7. Cleveland WS, Denby L, Liu C. Random location and scale effects: model building methods for a general class of models. *Computing Science and Statistics*. 2002; 32:3–10.
8. Piasecki TM, Trela CJ, Hedeker D, Mermelstein RJ. Smoking antecedents: separating between- and within-person effects of tobacco dependence in a multiwave ecological momentary assessment investigation of adolescent smoking. *Nicotine & Tobacco Research*. 2014; S:119–126.
9. Kapur K, Li X, Blood EA, Hedeker D. Bayesian mixed-effects location and scale models for multivariate longitudinal outcomes: an application to ecological momentary assessment data. *Statistics in Medicine*. 2015; 34(4):630–651. [PubMed: 25409923]
10. Dewey D, McDonald MK, Brown WJ, Boyd SJ, Bunnell BE, Schuldberg D. The impact of ecological momentary assessment on posttraumatic stress symptom trajectory. *Psychiatry Research*. 2015; 230(2):300–303. [PubMed: 26381184]
11. Rast P, Hofer SM, Sparks C. Modeling individual differences in within-person variation of negative and positive affect in a mixed effects location scale model using bugs/jags. *Multivariate Behavioral Research*. 2012; 47:177–200. [PubMed: 26734847]
12. Leckie G. Modeling heterogeneous variance-covariance components in two-level models. *Journal of Educational and Behavioral Statistics*. 2014; 39(5):307–332.
13. Goldstein H, Leckie G, Charlton C, Tilling K, Browne WJ. Multilevel growth curve models that incorporate a random coefficient model for the level 1 variance function. *Statistical Methods in Medical Research*. 2017 Epub: 962280217706728.
14. Bates D, Machler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. 2015; 67(1):1–48.
15. Bradley PC, Siddhartha C. Bayesian model choice via markov chain monte carlo methods. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1995; 57(3):473–484.
16. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell A. Stan: A probabilistic programming language. *Journal of Statistical Software*. 2017; 76(1):1–32.
17. McCulloch, CE., Neuhaus, JM. *Generalized Linear Mixed Models*. Wiley; New Jersey: 2005.
18. Hedeker, D., Gibbons, RD. *Longitudinal Data Analysis*. Wiley; New Jersey: 2006.
19. Fong Y, Rue H, Wakefield J. Bayesian inference for generalized linear mixed models. *Biostatistics*. 2010; 11(3):397–412. [PubMed: 19966070]
20. Gelman, A., Carlin, JB., Stern, HS., Dunson, DB., Vehtari, A., Rubin, DB. *Bayesian Data Analysis*. CRC Press; Florida: 2013.
21. Leonard T. A bayesian approach to the linear model with unequal variances. *Technometrics*. 1975; 17:95–102.
22. Hobert JP, Casella G. The effect of improper priors on gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*. 1996; 91(436):1461–1473.
23. Stan development team. Technical report. Bell Labs; 2014. Stan modeling language: User's guide and reference manual. URL <http://mc-stan.org>
24. Vehtari A, Gelman A, Gabry J. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*. 2017; 27(5):1413–1432.
25. Gwaltney CJ, Bartolomei R, Colby SM, Kahler CW. Ecological momentary assessment of adolescent smoking cessation. *Nicotine & Tobacco Research*. 2008; 10(7):1185–1190. [PubMed: 18629728]
26. Lukasiewicz M, Fareng M, Benyamina A, Blecha L, Reynaud M, Falissard B. Ecological momentary assessment in addiction. *Expert Review of Neurotherapeutics*. 2007; 7(8):939–950. [PubMed: 17678488]
27. Duncan AE. Hyperglycemia and perioperative glucose management. *Current Pharmaceutical Design*. 2012; 18(38):6195–6203. [PubMed: 22762467]
28. Hedeker D, Mermelstein RJ, Demirtas H. Modeling between- and within-subject variance in ecological momentary assessment (ema) data using mixed-effects location scale models. *Statistics in Medicine*. 2012; 31:3328–3336. [PubMed: 22419604]

29. Hedeker D, Demirtas H, Mermelstein RJ. A mixed ordinal location scale model for analysis of ecological momentary assessment (ema) data. *Statistics and Its Interface*. 2009; 2:391–402. [PubMed: 20357914]

Figure 1.

Mood assessments: erratic to consistent. (Random subject scale effect estimates are estimated to be $(-0.41, -3.64)$ and $(0.30, -0.25)$)

Figure 2.

Mood assessments: remains consistent or erratic. (Random subject scale effect estimates are estimated to be $(-1.09, -2.52)$ and $(1.36, 1.52)$)

Figure 3.
Visualization of the model mechanics

Table 1
Results from 100 simulations under the three level mixed effects location scale model : Mean Model Parameters

Model	$\beta^{intercept} = 1$			$\beta^{ubi} = 1$			$\beta^{wave} = 1$			$\beta^{bs} = 1$		
	Bias	AIW	COV	Bias	AIW	COV	Bias	AIW	COV	Bias	AIW	COV
Two Level MIRM HV	-0.0274	0.3973	46%	-0.0005	0.3953	91%	-0.0081	0.2075	17%	0.0003	0.0613	96%
Two Level MLS	-0.0292	0.4050	52%	0.0009	0.4000	94%	-0.0075	0.2103	19%	0.0011	0.0577	96%
Three Level MIRM HV	-0.0288	1.6993	98%	-0.0018	0.3976	95%	-0.0006	1.8039	97%	-0.0019	0.0510	98%
Three Level MLS	-0.0257	1.6873	99%	-0.0014	0.3989	93%	-0.0016	1.8150	95%	-0.0010	0.0404	94%

AIW: average 95% credible interval; COV: 95% coverage rate out of 100 simulations.

Table 2
Results from 100 simulations under the three level mixed effects location scale model : Variance Model Parameters

Model	$\sigma^{intercept} = 0.3$			$\sigma^{subj} = 0.2$			$\sigma^{var} = 0.1$		
	Bias	AIW	COV	Bias	AIW	COV	Bias	AIW	COV
Two Level MIRM HV	0.6109	0.0608	1	-0.0609	0.0567	19	-0.0096	0.0712	16
Two Level MLS	0.5350	0.1499	5	-0.0730	0.1439	49	0.0071	0.1106	17
Three Level MIRM HV	0.2410	0.0605	5	-0.0004	0.0570	47	-0.0402	0.0672	12
Three Level MLS	-0.0038	0.8596	97	0.0040	0.2081	95	-0.0155	0.9114	99

AIW: average 95% credible interval; COV: 95% coverage rate out of 100 simulations.

Table 3
Application of the Bayesian Three Level Mixed Effects Location Scale Model on Adolescent Smoking Study

Parameters	MRM			MRM with HV			MLS		
	Estimate	95% Credible Interval		Estimate	95% Credible Interval		Estimate	95% Credible Interval	
$\beta_{intercept}$	6.6678	(6.3530, 7.0017)		6.6533	(6.3174, 6.9810)		6.6739	(6.3507, 6.9960)	
β_{male}	-0.0409	(-0.3158, 0.2401)		-0.0415	(-0.3353, 0.2464)		-0.0455	(-0.3166, 0.2121)	
$\beta_{smksubj}$	0.5538	(-0.5657, 1.6776)		0.5886	(-0.5824, 1.7126)		0.7113	(-0.4083, 1.8347)	
$\beta_{smkwave}$	0.0327	(-0.6442, 0.7056)		0.0278	(-0.6352, 0.6695)		0.0385	(-0.6297, 0.7041)	
β_{smkobs}	0.2310	(0.1927, 0.2685)		0.2260	(0.1896, 0.2615)		0.0965	(0.0689, 0.1238)	
β_{wave}	0.4498	(0.2594, 0.6428)		0.4465	(0.2594, 0.6445)		0.4463	(0.2523, 0.6411)	
$\alpha_{intercept}$	0.6850	(0.6700, 0.6997)		1.0930	(1.0513, 1.1355)		0.8728	(0.6868, 1.0700)	
α_{male}	—	—		-0.1693	(-0.2007, -0.1375)		-0.1379	(-0.3180, 0.0270)	
$\alpha_{smksubj}$	—	—		-0.7555	(-0.8816, -0.6342)		-0.5698	(-1.2555, 0.1118)	
$\alpha_{smkwave}$	—	—		-0.3535	(-0.5465, -0.1635)		-0.0578	(-0.5989, 0.4788)	
α_{smkobs}	—	—		-0.0666	(-0.1043, -0.0284)		-0.0600	(-0.1030, -0.0192)	
α_{wave}	—	—		-0.2060	(-0.2380, -0.1734)		-0.3211	(-0.4645, -0.1800)	
location: $\sigma^2_{subj int}$	0.9808	(0.7293, 1.2807)		0.9689	(0.7087, 1.2978)		1.0020	(0.7583, 1.3068)	
location: $\sigma^2_{subj slope}$	0.6393	(0.3389, 1.0178)		0.6449	(0.3106, 1.1147)		0.6902	(0.3711, 1.0580)	
location : $\sigma_{int slope}$	-0.1205	(-0.3740, 0.0930)		-0.1171	(-0.4175, 0.1108)		-0.1434	(-0.4051, 0.0698)	
location: σ^2_{wave}	0.4080	(0.3597, 0.5058)		0.4582	(0.3563, 0.4919)		0.3794	(0.3201, 0.4505)	
scale: $\sigma^2_{subj int}$	—	—		—	—		0.2755	(0.1785, 0.3935)	
scale: $\sigma^2_{subj slope}$	—	—		—	—		0.4298	(0.2518, 0.6570)	
scale : $\sigma_{int slope}$	—	—		—	—		-0.0953	(-0.2263, 0.0135)	

Parameters	MRM		MRM with HV		MLS	
	Estimate	95% Credible Interval	Estimate	95% Credible Interval	Estimate	95% Credible Interval
$scale: \sigma_{wave}^2$	—	—	—	—	0.2729	(0.2218, 0.3138)
$elpd_{100}$	−57755.1 (182.8)		−57532.3 (184.1)		−53721.5 (204.7)	