



UHASSELT



Maastricht University

KNOWLEDGE IN ACTION

Faculty of Sciences
School for Information Technology

Master of Statistics

Masterthesis

Bayesian hierarchical models for powder manufacturing

Vivienn Weru

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics, specialization Biostatistics

SUPERVISOR :

Prof. dr. Ziv SHKEDY

SUPERVISOR :

Dr. Martin OTAVA

Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.



UHASSELT

KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

2017
2018



Maastricht University

Faculty of Sciences

School for Information Technology

Master of Statistics

Masterthesis

Bayesian hierarchical models for powder manufacturing

Vivienn Weru

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics, specialization Biostatistics

SUPERVISOR :

Prof. dr. Ziv SHKEDY

SUPERVISOR :

Dr. Martin OTAVA

Contents

| | | |
|----------|--------------------------------------------------------------|-----------|
| 1 | Introduction | 1 |
| 2 | Data | 3 |
| 2.1 | Structure | 3 |
| 2.2 | Exploratory analysis | 4 |
| 3 | Methodology | 7 |
| 3.1 | The model | 7 |
| 3.2 | Implementation of the model | 7 |
| 3.3 | Bayesian Regression Models using Stan (BRMS) | 8 |
| 3.4 | Model diagnostics | 9 |
| 3.4.1 | Checking convergence | 9 |
| 3.4.2 | Posterior predictive checks | 10 |
| 3.4.3 | Predictive approach to outlier detection | 10 |
| 3.5 | Posterior prediction | 11 |
| 3.6 | Software | 11 |
| 4 | Results | 13 |
| 4.1 | Primary Results | 13 |
| 4.1.1 | Results from R2jags | 13 |
| 4.2 | Model diagnostics-R2jags | 14 |
| 4.2.1 | Sensitivity to priors | 14 |
| 4.2.2 | Convergence of the R2jags models | 15 |
| 4.2.3 | Posterior predictive checks | 17 |
| 4.3 | Variance Partition Coefficient | 18 |
| 4.4 | Prediction of future batches | 19 |
| 4.5 | Secondary results | 20 |
| 4.5.1 | Results from frequentist and Brms on complete data | 20 |
| 5 | Discussion | 23 |
| 6 | Limitation | 25 |
| 7 | Conclusion and recommendation | 27 |

List of Figures

| | | |
|---|-----------------------------------------------|----|
| 1 | Data hierarchy structure | 4 |
| 2 | Exploratory plots | 5 |
| 3 | Trace plots | 15 |
| 4 | Autocorrelation plots | 16 |
| 5 | Running-mean plot | 16 |
| 6 | Posterior Predictive Ordinate plots | 18 |
| 7 | Index plot of 1/CPO | 31 |
| 8 | Trace plot for the BRMS parameters | 32 |

List of Tables

| | | |
|----|-----------------------------------------------------------------|----|
| 1 | Variables description | 3 |
| 2 | Descriptive statistics | 5 |
| 3 | Parameter estimates-Complete data | 13 |
| 4 | Parameter estimates-Incomplete data | 13 |
| 5 | Different prior for the "fixed effects" | 14 |
| 6 | Different priors for the variance components | 15 |
| 7 | Geweke Z-values | 17 |
| 8 | MC error estimated via a time-series approach | 17 |
| 9 | Assessing normality of random effects | 18 |
| 10 | Variance Partition Coefficient | 19 |
| 11 | Proportion that passed acceptance criteria | 19 |
| 12 | Frequentist and Brms for the complete data | 21 |
| 13 | Parameter estimates for BRMS for the complete dataset | 31 |

Acknowledgement

Special thanks to Prof. dr. Ziv Shkedy (Hasselt University) and Dr. Martin Otava (Janssen: Pharmaceutical Companies of Johnson & Johnson) for their guidance and motivation during the thesis period and their willingness to respond to my questions.

My profound gratitude also goes to my mom and grandma for their unwavering encouragement and support all through my study years. To my friends and classmates, I truly thank you for your support in the two years.

I acknowledge the financial assistance from Vlaamse Inter Universitaire Raad (VLIR) who made my dream of attaining a Master degree in Biostatistics a reality. To God Almighty, thank you for granting me strength throughout.

Vivienn Weru,
Diepenbeek, June 15, 2018.

Abstract

Background: Tablets and capsules are the most widely used form of drug delivery. As such, pharmaceutical manufacturing of powdered ingredients prior to tablet/capsule manufacturing requires analysis of the data with several levels of hierarchy. Next to different batches of the product, different bags of resulting material are available and samples are taken within multiple locations across the bags. The evaluation of the quality of manufacturing process requires the characterization of the mean structure, but also careful evaluation of the sources of variability.

Objectives: To investigate effect of location in the bag from where a sample of the product is obtained and to characterize the total variability distribution across possible variance components. In addition, to predict behaviour of future batches.

Methodology: A Bayesian linear mixed effects model was fitted. This model was fitted using R2jags and Brms packages in R. A frequentist equivalent was also fit for comparison.

Results: Exploratory analysis showed that samples obtained at the beginning of the bag had the least percentage of Active Pharmaceutical Ingredient (API). Location was in general found to have an effect on the percentage of API. Several sources of variation were distinguished; variation across batches, variation across bags (within batches), variation across analytical runs and residual variability. Of these, variation across batches explained much of the total variability. In terms of prediction, new samples taken from existing batches and measured using existing analytical runs would almost always pass the acceptance criteria. Future batches would also almost always pass the acceptance criteria unless produced with low quality raw materials.

Conclusion: R2jags was shown to provide a more flexible approach as compared to the Brms. Compared to frequentist, Brms was shown to be a better approach for a balanced design dataset.

Key words: Bayesian modeling; Hierarchical model; Prediction; Brms; R2jags; Variation partition.

1 Introduction

Tablets and capsules are the most widely used form of drug delivery and therefore drug powder production needs to be monitored in all ways. Every pharmaceutical company's aim is to produce drugs that are stable with excellent dose uniformity, have an increased shelf-life and that are well-accepted by users. Other than increasing the shelf-life and stability, another main aim of powder drug manufacturing is to reduce the weight and to develop an easier mode of drug delivery. This development is challenging since every step to simplicity of a drug is less than straightforward. Pharmacologically active substances used in drug production include natural products and synthetic drugs with emphasis on the latter due to advancement in chemical and computer technologies [1]. In the manufacturing process, the active ingredients are mixed with other pharmaceutical excipients that have limited or no pharmaceutical value to provide the pharmacological properties needed to form the final drug. The actual drug production involves fermentation, chemical synthesis and biological extraction which can be done as a discrete batch and/or a continuous process [2]. In drug production, manufacturing occurs by first manufacturing the active pharmaceutical ingredient (API) and then by mixing the API with excipients from where the final drugs are prepared. API is the active part of the drug containing the medicinal part. This part of the production is the most involving and consequently the most expensive part of production [3]. High quality API are therefore expected to shield end-users from harm.

After the production, the next stage involves converting the product to an administrative-friendly form, for example to a powder [2]. The process of producing dry powder from liquid suspensions is called spray-drying [4]. This process works by using high temperature gas to evaporate out the liquid to form particles. The process starts with a feed preparation of the suspension or paste that is free of impurities and contains the Active Pharmaceutical Ingredient (API) and other excipients. The next process turns this paste into droplets in a process called atomization. The droplets are then passed through a nozzle into a chamber where hot gas is being blown into. As a result, the droplets lose their moisture content and becomes powder. Powder characteristics that can be influenced by the spray-drying process include density, particle size distribution, retention of volatile components, mechanical stability as well as storage stability [5].

Another characteristic that can be affected by spray-drying is the purity of the active ingredient. Spray drying can increase stability of the API which in turn improves its solubility and bioavailability [6]. This, however, will only be achieved if the spray-drying is regulated at appropriate parameters. The European Agency for the Evaluation of Medicinal Products, Human Medicines Evaluation (EMA) Unit puts weight on establishing guidelines on activities that may affect the active pharmaceutical ingredient. Therefore, spray-drying characteristics such as temperature, drying time, drying system among others should be keenly monitored and the appropriate values should be established to ensure homogeneous results. A study by [7] revealed that the amount of API in a sample was related to the spray drying of the feed concentration.

Some pharmaceutical companies operate in international markets subjecting their activities, especially in drug development, to strict regulations [8]. Quality issues in the pharmaceutical industry are regulated by several organizations, including the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), Food and Drug Administration (FDA) and EMEA. This is done primarily to protect the end users of the products by ensuring the pharmaceutical industry upholds good ethics [9]. Evaluating the manufacturing process, therefore, enables detection of any possible issues that may be occurring.

Variation in a manufacturing process influences the quality of manufactured products, thus the evaluation of the variability is vital [10]. This may require analysis of the data on several levels of hierarchy. In the case considered in this report, several sources of variability were distinguished; variability across batches, variability across bags in the batches, variability across measurements and variability across analytical runs. Of the various sources of variability, batch to batch variability is usually the highest and cannot be ignored since by so doing, the process quality may be compromised. Being able to quantify and model this kind of variation is vital in ensuring that the process does not fall short of manufacturing qualifications [11]. Interest of this study was to determine the percentage of total variability attributed to the above sources of variation and to investigate whether there was an effect of location. Further, predicting behaviour of future batches on whether they would be acceptable.

2 Data

2.1 Structure

There were two datasets used both of which were simulated; a complete design dataset and an incomplete design dataset. The response is the percentage of the Active Pharmaceutical Ingredient (API) in a sample. Table 1 below lists the variables in the data sets.

Table 1: Variables description

| Variable | Type | Description |
|----------------|--------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Assay | Continuous (Response) | Indicates the % of amount of API in the sample |
| Location | Categorical (4 levels: Beginning, Middle, End and Composite) | Indicates location in the bag from where the sample was obtained. A composite sample was taken as the equal mixture of samples from the other three locations. |
| Bag | Categorical | Indicates the bag in the batch where the assay was obtained |
| Batch | Categorical | Indicates the batch from which the assay was obtained |
| Analytical run | Categorical | Indicates the measurement system that was used to measure the assay after the sample was obtained |

There is a hierarchy feature to this data such that there are several measurements obtained in bags which are nested within batch. The model therefore differentiates several sources of variability in the response as due to; variation across batches, variation across bags (within batches), and variation across the measurements (within bags nested within batches). Another source of variation in the response, that does not fall in the above hierarchy, is attributed to the analytical run that was used to measure the assay with several samples per run measured together. These samples are considered as clustered, because same analytical run is prepared by same analyst and same preparation of samples is done. Figure 1 shows the hierarchical structure of the data.

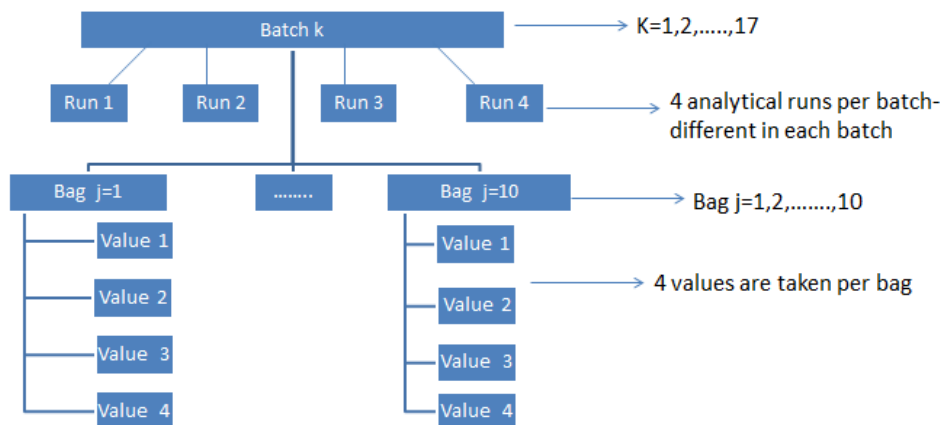


Figure 1: Data hierarchy structure

The complete design implies that 4 samples were obtained from each of the 10 bags in each of the 17 batches and 10 samples per run were measured. In total, 680 assay measurements were obtained for the complete design. The incomplete design implies that for some bags, not all 4 samples were obtained and not all 10 bags were evaluated in each of the 17 batches. The structure of the analytical run is also different with less than 10 samples measured for some runs. In total, 182 assay measurements were obtained for the incomplete design. The two datasets are used to provide more insights on the parameters change in presence of a balanced and unbalanced design. For the rest of this report these datasets will, for simplicity, be referred to as the complete and incomplete data set.

2.2 Exploratory analysis

The most important exploratory analysis was to visually evaluate whether there was a time trend in the production of the batches. Figure 2 shows this evaluation for the complete and the incomplete data set. The batches were produced such that Batch 1 was the first to be produced while Batch 17 was the last. There, however, does not seem to be any definite time trend. For both data sets, Assay obtained from the Beginning of the bag tends to have the lowest percentage of the active pharmaceutical ingredient. From Figure 2, Batch 4 seems to have the lowest percentage of API on average.

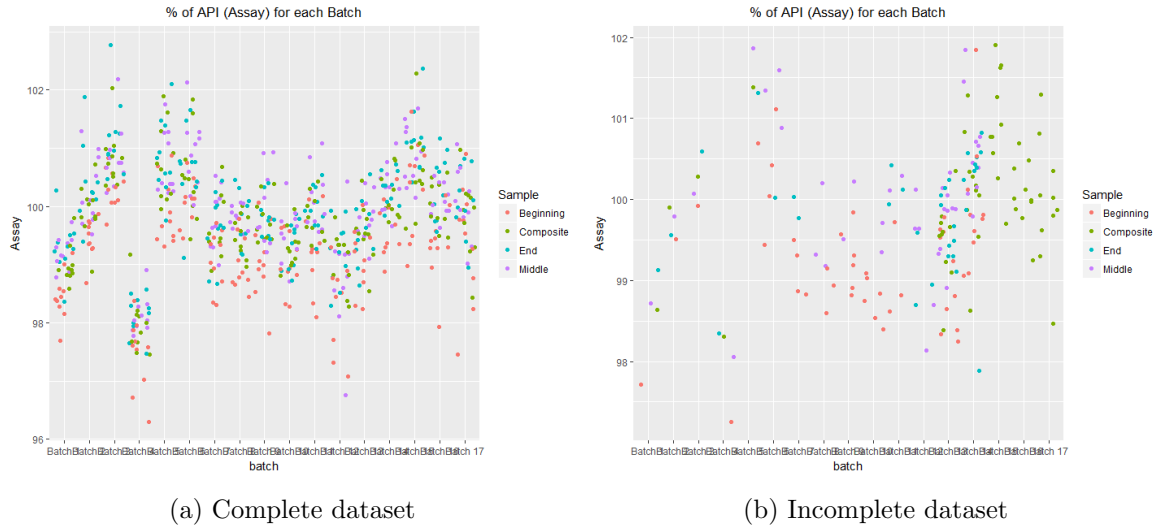


Figure 2: Exploratory plots

Table 2 displays the minimum and maximum assay for each of the datasets.

Table 2: Descriptive statistics

| | Min Assay | Max Assay |
|-----------------|-----------|-----------|
| Complete data | 96.30 | 102.77 |
| Incomplete data | 97.25 | 101.91 |

The minimum and maximum assay for the complete dataset is 96.29% and 102.77%, respectively. For the incomplete dataset, the minimum and maximum assay are 97.25% and 101.91%, respectively.



3 Methodology

3.1 The model

A linear mixed model was fitted to the data with the model described as;

$$Y_{ijkm} = \alpha_0 + \beta_1 Location_{ijkm} + b_{0k} + b_{0j(k)} + b_{0m} + \epsilon_{ijkm},$$

$i = 1, \dots, 680$; $j = 1, \dots, 10$; $k = 1, \dots, 17$; $m = 1, \dots, 68$ for the balanced design.

and:

Y_{ijkm} =the i^{th} measurement of assay of the j^{th} bag in the k^{th} batch and obtained in the m^{th} analytical run.

α_0 = process mean

$Location_{ijkm}$ =location in the bag where the sample was obtained: beginning, medium, end or composite.

b_{0k} =random effect of the k^{th} batch: $b_{0k} \sim N(0, \sigma_{b_{0k}}^2)$

$b_{0j(k)}$ =random effect of j^{th} bag of k^{th} batch: $b_{0j(k)} \sim N(0, \sigma_{b_{0j(k)}}^2)$

b_{0m} =random effect of the m^{th} analytical run: $b_{0m} \sim N(0, \sigma_{b_{0m}}^2)$

ϵ_{ijkm} = residual error: $\epsilon_{ijkm} \sim N(0, \sigma_\epsilon^2)$ for sample obtained at the beginning, middle, end and $\epsilon_{ijkm} \sim N(0, \frac{\sigma_\epsilon^2}{3})$ for sample obtained as composite.

Non-informative priors were applied to all the model parameters.

Priors for α_0 and β_i :

$$\begin{aligned}\alpha_0 &\sim N(0, \sigma_\alpha^2) \\ \beta_i &\sim N(0, \sigma_\beta^2)\end{aligned}$$

Priors for b_{0k} , $b_{0j(k)}$, b_{0m} :

$$\begin{aligned}b_{0k} &\sim N(0, \sigma_{b_{0k}}^2) \\ b_{0j(k)} &\sim N(0, \sigma_{b_{0j(k)}}^2) \\ b_{0m} &\sim N(0, \sigma_{b_{0m}}^2)\end{aligned}$$

Priors for the variance components:

$$\begin{aligned}\sigma_{res}^2 &\sim IG(\epsilon, \epsilon) \\ \sigma_{b's} &\sim U(0, c)\end{aligned}$$

In fitting the model, a frequentist as well as a Bayesian approach were considered.

3.2 Implementation of the model

In a hierarchical model, there are parameters that describe the individual's behaviour with the distribution of individual parameters within a group described by higher-level distribution. The data from different individuals within a group are usually dependent and therefore the estimation of the group-level parameters is done better by the joint individual-level parameters [12]. The Bayesian framework allows use of prior knowledge in estimation of parameters. This prior knowledge is combined with the current data to yield the posterior

distribution where summary measures for the model parameters are obtained [13]. Usually analytical approach in estimating the posterior distribution for model parameters is not available. To sample from such posterior distributions, the so-called Markov Chain Monte-Carlo (MCMC) is used.

MCMC is a method that aids sampling from a probability distribution combining the property of Monte-Carlo that allows evaluating the properties of a distribution by investigating random samples from the distribution and Markov chain which indicates that the random samples are generated in a sequential process. The Markov chain, however, imply that given θ a vector of parameters and k iterations and given θ^k , θ^{k+1} does not depend on θ^{k-1} , θ^{k-2} and so on [14]. Common MCMC procedures include Gibbs sampling and the so-called Metropolis(-Hastings)[15]. The Gibbs sampling operates on the property that a multivariate distribution is determined by its conditional distributions called the full conditionals. In other words, for parameters $\theta_1, \theta_2, \dots, \theta_d$, the full conditional distributions are given by $p(\theta_j | \theta_1^k, \theta_2^k, \dots, \theta_{(j-1)}^k, \theta_{(j+1)}^k, \dots, \theta_d^k, \mathbf{y})$ implying that θ_j is conditioned on all other parameters. More on MCMC can be found in [14].

In fitting the model, several computational considerations are made. An example is to include a burn-in portion. A burn-in portion is included for convenience to help in reducing the influence of initial values on the posterior inference. This is required since there is no guarantee that the initial samples are drawn from the target distribution. To prevent an increase in the Monte Carlo error and to ensure sufficient approximation of the distribution, the remaining part is usually made longer. Usually, using better starting values can eliminate this problem but good initial values may be difficult to find [14].

Thinning can also be performed whereby instead of saving all iterations, we only save the n^{th} iteration, with $n > 1$. Thinning works by recognizing that the many iterations are highly correlated and there is no benefit in saving all the iterations. Instead, we save only the n^{th} iteration. Starting values for the model parameters were obtained by random generation from respective random distributions. Using random distributions to select the initial values from, ascertains that different chains start at different points which helps in attaining convergence.

3.3 Bayesian Regression Models using Stan (BRMS)

MCMC has been the main algorithm in sampling from the posterior distribution implemented in the Jags/BUGS family. The development of such a method has contributed to the increased use of Bayesian inference. Nonetheless, some models take longer to converge. To aid in this, an algorithm that increases efficiency and leads to faster convergence has been proposed [16]. This is the so-called Hamiltonian Monte Carlo (HMC). The original HMC requires experts to implement which explains why the algorithm has not been popular. To remedy this, an automatic HMC algorithm called a no-U-turn sampler (NUTS) has been created and packaged into what is now known as Stan software [17]. Bayesian Regres-

sion Models using Stan (BRMS) is, therefore, a package that connects R software to Stan programming. Our interest was to also investigate the benefits of Brms to other Bayesian packages and to provide guidance to the pharmaceutical industry in terms of options in Bayesian inference.

3.4 Model diagnostics

3.4.1 Checking convergence

Unlike in an analytical approach where convergence is never a problem, in a numerical approach convergence could be a big problem and therefore it has to be monitored. For an MCMC algorithm, the aim is to investigate how close we are to the true posterior distribution. Two features in checking convergence include; checking stationarity of the chain by evaluating the burn-in part of the Markov chain and determining the accuracy of the posterior summary measures. Common graphical approaches to checking convergence, mixing rate and stability include the trace plot, autocorrelation plot and running-mean plot, respectively. A trace plot is produced for each parameter as well as for the parameter vector to monitor the Markov chain in a joint form. To monitor the Markov chain jointly, a trace plot of the log of the likelihood or of the posterior density is used. The autocorrelation plot is used to check the mixing rate of a chain while a running-mean plot is used to explore stability of the mean of the sampled values [14].

The above graphical checks are important but they may be subjective, implying that formal diagnostics should also be used. A more formal diagnostic is the Geweke diagnostic that evaluates the size of the burn-in part for a single chain. The diagnostic makes use of a frequentist significance test by comparing the means of an early (E) and a late part (L) of the chain. The posterior means of the two parts, indicated as $\bar{\theta}_E$ and $\bar{\theta}_L$ are then compared with a Z-test given by; $\frac{\bar{\theta}_E - \bar{\theta}_L}{\sqrt{s_E^2/n_E + s_L^2/n_L}}$. A significant Z-test implies that the burn-in part was not taken enough or the total chain was not long enough or both [14].

Another quantity used to assess convergence is the effective number of simulation draws. The effective number of simulation draws is smaller since the Markov chain simulations are autocorrelated. A converged model should have this quantity to be at least 100 for each of the parameters. Brooks-Gelman-Rubin (BGR) diagnostic is another formal diagnostic tool for detecting non-stationarity that, unlike the above diagnostics, evaluates differences between multiple Markov chains having different starting positions. This approach is based on the so-called potential scale reduction factor (\hat{R}). The ratio is the square root of the mixture variance obtained for the chains mixed together divided by the average within-chain variance obtained from averaging the variance of each of the parameters in each chain. With the chains well mixed, convergence would be achieved and the ratio above would equal 1, indicating equivalence of the distribution of simulations between and within chains. Using multiple chains can also be used to check problems with convergence where, similarly to the above mentioned, differences in the distribution of samples from different chains may indicate convergence problems [18].

3.4.2 Posterior predictive checks

To check the fit of the model, we check for differences between the sampled data from the model and the observed data which can be done through posterior predictive checks in a Bayesian context. Considering the data, \mathbf{y} , some test statistics include $T_{min} = \min(\mathbf{y})$ and $T_{max} = \max(\mathbf{y})$, from which a posterior predictive p-value can be determined by;

$$p_T = \int p_C(\mathbf{y}, \theta) p(\theta | \mathbf{y}) d\theta$$

where $p_C(\mathbf{y}, \theta)$ is the sampling distribution of the test statistic, $T(\mathbf{y})$, θ is a vector of parameters and $p(\theta | \mathbf{y})$ is the posterior distribution. In a nutshell, let $T(y)$ be a test statistic based on the data and $T(y^{rep})$ be a test statistic based on replicated data from the fitted model. The Bayesian p-value is then given by $pr[T(y^{rep}, \theta) \geq T(y, \theta | y)]$ and denotes the proportion of replicated values for the test statistic that are equal or more than the values of the test statistic based on the observed data [19]. Values of this p-value that are small (≤ 0.05) indicate possible lack of fit.

The posterior predictive check can also be used to check distributional assumptions based on the discrepancy measure, $D(y, \theta)$, which extends the above classical test statistics allowing it to depend on the parameters as well. A posterior predictive p-value based on this approach is then computed with small values indicating issues with normality. These discrepancy measures can be based on skewness and kurtosis. For hierarchical models, the several levels of hierarchy call for the use of additional discrepancy measures that are more inclined to their structure. An example is the Kolmogorov-Smirnov discrepancy measure suggested by [20], defined as; $D_{KS}(\mathbf{y}, \theta) = \max_{i \in \{1, \dots, n\}} [\phi(y_i | \mu, \sigma) - (i-1)/n, i/n - \phi(y_i | \mu, \sigma)]$. Another approach was suggested by [21] to check for distributional assumptions of the random effects in a normal hierarchical model and is defined as; $D_{SS}(\mathbf{y}, \theta, \psi) = |\theta_{max} - \theta_{med}| - |\theta_{min} - \theta_{med}|$.

3.4.3 Predictive approach to outlier detection

It is also important to check for outlying observations, which can be done in a Bayesian context using Bayesian residual analysis, predictive approaches and using DIC diagnostics. Here, predictive approach to outlier detection is adopted which uses the so-called posterior predictive ordinate (PPO_i), defined as;

$$PPO_i = p(y_i | \mathbf{y}) = \int p(y_i | \theta) p(\theta | \mathbf{y}) d\theta$$

It is the value of the posterior predictive distribution (PPD) evaluated at y_i . A very low value of the PPO for the i^{th} observation could imply that this observation is outlying. This approach, however, uses the observed data, \mathbf{y} , twice and a better approach would be one that evaluates the PPD in y_i using the posterior density based on a sample without y_i . Such an approach that uses cross-validation is the so-called conditional posterior ordinate (CPO) and is defined as;

$$CPO_i = p(y_i | \mathbf{y}_{(i)}) = \int p(y_i | \theta) p(\theta | \mathbf{y}_{(i)}) d\theta$$

In other words, CPO_i represents the posterior probability of observing y_i when the model is fit to all data except y_i . More details on outlier detection can be found in [14].

3.5 Posterior prediction

In pharmaceutical drug production, it is of interest to predict whether batches produced in the future will be acceptable depending on what we know now. This prediction can therefore be done taking into account the location in the bag where assay sample are taken, the bag, the batch as well as the analytical run that was used to measure the assay. The Bayesian framework is applied to obtain individual predictions while considering the uncertainty within the parameter estimates. This was also another reason why a Bayesian approach was adopted since classical approaches to prediction may fail to take all uncertainty about model parameters into account. In the prediction, some terms in the model can be examined at conditional or marginal values. First, prediction for new samples was done conditional on all the random effects taking the process mean to be equal to 100%. 100% being the ideal percentage of API in the sample. Second, prediction was done for a given value of the batch mean to evaluate behavior of new samples when the batch mean is shifted.

3.6 Software

The models were implemented in R, specifically libraries connecting Bayesian modeling to R. The frequentist approach was implemented in lme4, the Bayesian approach was implemented in both R2jags, and Brms. Brms connects R with Stan programming while R2jags connects R to Jags programming.

4 Results

4.1 Primary Results

4.1.1 Results from R2jags

The composite sample was taken as an equal mixture of samples from the other three locations. This implies that a sample obtained from a composite location had inherently different residual variability, being the average of the three other samples. The Bayesian approach in jags provided a flexible approach to incorporate this information. Thus, in the model specification in the R2jags, we define the residual variability for each response depending on where in the bag that response was taken. The jags model was fit with a $N(0,10E-4)$ prior for the "fixed effects" parameters (which are random in Bayesian framework) and $U(0,100)$ for the variance parameters. The results using these priors are displayed in Table 3 for the complete dataset.

Table 3: Parameter estimates-Complete data

| | mean | sd | 2.50% | 97.5% | Rhat | n.eff |
|------------------------|--------|-------|--------|--------|------|-------|
| α | 99.26 | 0.20 | 98.86 | 99.66 | 1.00 | 5400 |
| β_1 | 0.53 | 0.04 | 0.44 | 0.61 | 1.00 | 5400 |
| β_2 | 0.71 | 0.05 | 0.61 | 0.80 | 1.00 | 3100 |
| β_3 | 0.72 | 0.05 | 0.63 | 0.82 | 1.00 | 5400 |
| deviance | 589.85 | 25.84 | 540.50 | 641.14 | 1.00 | 5400 |
| Sources of variability | | | | | | |
| σ_{res} | 0.43 | 0.01 | 0.40 | 0.46 | 1.00 | 5400 |
| σ_{batch} | 0.78 | 0.16 | 0.54 | 1.17 | 1.00 | 4200 |
| $\sigma_{bags(batch)}$ | 0.40 | 0.03 | 0.35 | 0.46 | 1.00 | 5400 |
| $\sigma_{A.run}$ | 0.07 | 0.04 | 0.006 | 0.14 | 1.02 | 400 |

Using the same prior on the incomplete dataset had the results displayed in Table 4.

Table 4: Parameter estimates-Incomplete data

| | mean | sd | 2.50% | 97.5% | Rhat | n.eff |
|------------------------|--------|-------|--------|--------|------|-------|
| α | 99.33 | 0.21 | 98.90 | 99.74 | 1.00 | 5400 |
| β_1 | 0.50 | 0.15 | 0.24 | 0.81 | 1.00 | 2000 |
| β_2 | 0.73 | 0.11 | 0.52 | 0.96 | 1.00 | 5400 |
| β_3 | 0.54 | 0.12 | 0.30 | 0.77 | 1.00 | 5400 |
| deviance | 160.42 | 22.36 | 119.56 | 205.80 | 1.00 | 2900 |
| Sources of variability | | | | | | |
| σ_{res} | 0.45 | 0.04 | 0.38 | 0.53 | 1.00 | 5400 |
| σ_{batch} | 0.70 | 0.19 | 0.38 | 1.15 | 1.01 | 5300 |
| $\sigma_{bags(batch)}$ | 0.48 | 0.05 | 0.39 | 0.59 | 1.00 | 5400 |
| $\sigma_{A.run}$ | 0.16 | 0.10 | 0.01 | 0.39 | 1.02 | 550 |

From these results, there was a significant effect of location on the assay for both datasets. The jags model fitted to the incomplete data was the primary model of choice and interpretation is done with respect to this model. This model was chosen mainly due to its flexibility and was fitted to the incomplete design dataset that reflects more of what is to be expected in practice. From Table 4 above, β_1 is the parameter estimate corresponding to samples obtained as composite, β_2 is the parameter estimate corresponding to samples obtained at the middle while β_3 corresponds to samples obtained at the end of the bag. The reference category was beginning category. The 95% credible interval (CI) does not include zero, implying a significant effect of location on the assay. Samples obtained from other locations in the bag were found to have higher percentage of API as compared to samples obtained at the beginning of the bag, with middle location having the highest. The proportion of variability attributed to each source of variation is shown in Table 9 (section 4.3).

4.2 Model diagnostics-R2jags

4.2.1 Sensitivity to priors

The posterior distribution might be influenced by the prior, making the choice of prior vital. Several priors were therefore explored. In choosing the priors, only the complete data set was used to take advantage of the benefit of a complete design. Sensitivity of the "fixed effects" parameters to changes in the prior was evaluated. This was done by changing the prior from $N(0,10E-4)$ to $N(0,10E-5)$ and to $N(0,10E-6)$. The low precision (high variance) defines a wide range for the coefficients implying that the prior information is contributing very little information to the inference about these coefficients [22]. The results are shown in Table 5, posterior mean (posterior std. dev.). From this table, varying the prior had no relevant impact on the parameters.

Table 5: Different prior for the "fixed effects"

| Parameter | $N(0,10E-4)$ | $N(0,10E-5)$ | $N(0,10E-6)$ |
|-----------|--------------|--------------|--------------|
| α | 99.26 (0.20) | 99.25 (0.20) | 99.25 (0.20) |
| β_1 | 0.53 (0.04) | 0.53 (0.04) | 0.53 (0.04) |
| β_2 | 0.71 (0.05) | 0.71 (0.05) | 0.70 (0.05) |
| β_3 | 0.72 (0.05) | 0.72 (0.05) | 0.72 (0.05) |

Several priors were suggested for the variance of the random effects; $U(0,100)$, $U(0,1000)$ and $IG(\epsilon, \epsilon)$, with $\epsilon = 0.001, 0.0001$ for the residual variance. A choice was made between $U(0, c)$ with $c=100, 1000$ for $\sigma_{b's}$ and the inverse-gamma prior for σ_{res}^2 , $IG(\epsilon, \epsilon)$, with ϵ chosen to be very small. For the inverse gamma prior on the higher level variances for a hierarchical model, the choice of ϵ has an impact on the posterior distribution making it very dependent on the prior parameter values [23]. As a result, only the uniform prior was considered for the higher level variances while both uniform and inverse gamma were considered for the level-1 variance. To make a choice on these priors, parameters for the random components were all explored through posterior summary measures. Table 6 below shows the comparison of

these priors. For each prior, three chains were initiated with 100,000 iterations, a thin of 50 and a burn-in of 10,000. Convergence for all was achieved almost immediately. Sensitivity of the parameters to the change in priors was explored.

Table 6: Different priors for the variance components

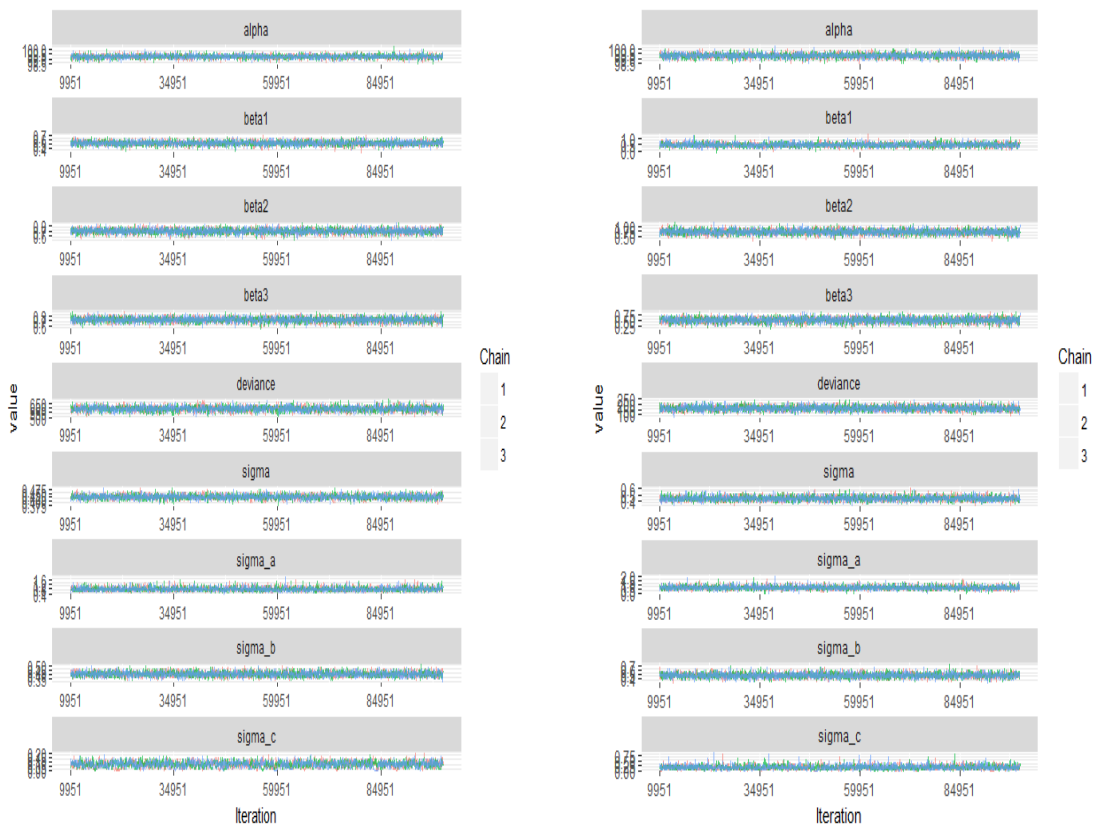
| | U(0,100) | U(0,1000) | IG(0.001,0.001) | IG(0.0001,0.0001) |
|-----------------------|----------|-----------|-----------------|-------------------|
| σ_{res} | 0.43 | 0.43 | 0.43 | 0.43 |
| σ_{batch} | 0.78 | 0.78 | - | - |
| $\sigma_{bag(batch)}$ | 0.40 | 0.40 | - | - |
| $\sigma_{A.run}$ | 0.07 | 0.07 | - | - |

”-” means that the prior was not used for that parameter

The results obtained for the variance parameters remained consistent over the range of the priors indicating that the choice of prior had little impact on the posterior estimates.

4.2.2 Convergence of the R2jags models

When using the MCMC approach, convergence has to be monitored keenly. Trace plots of the parameters are shown in Figure 3 giving an indication of achieved stationarity.



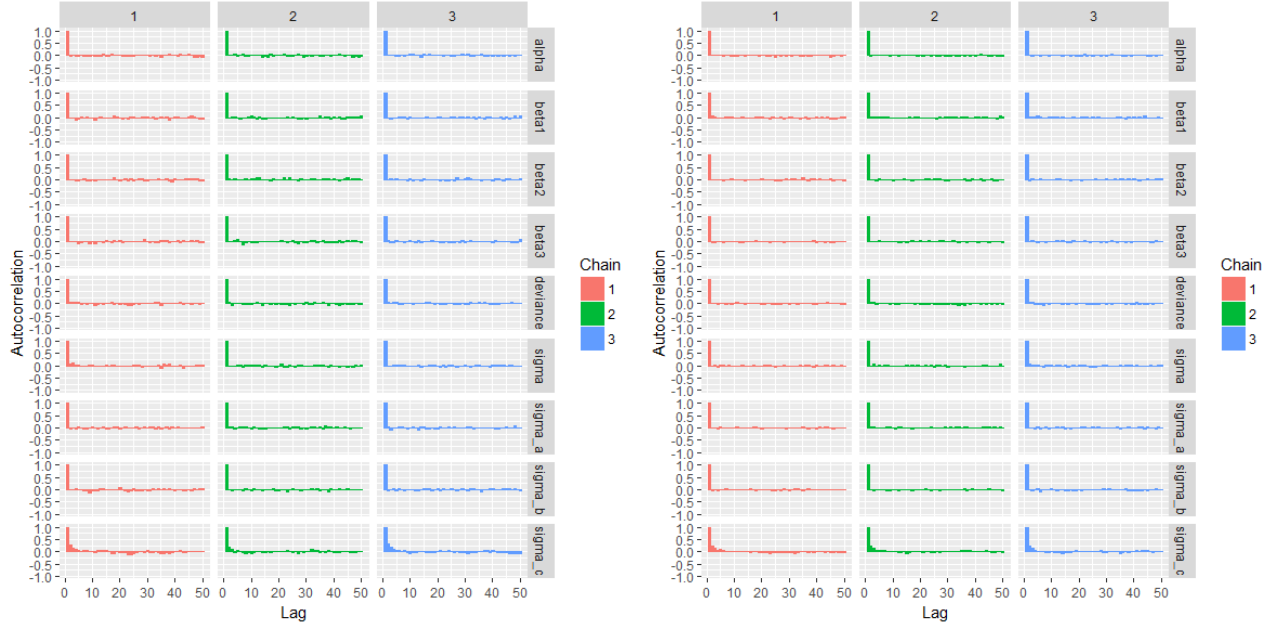
(a) Trace plots-complete data

(b) Trace plots-incomplete data

Figure 3: Trace plots

Autocorrelation plots were also obtained as shown in Figure 4. Though this plot is not

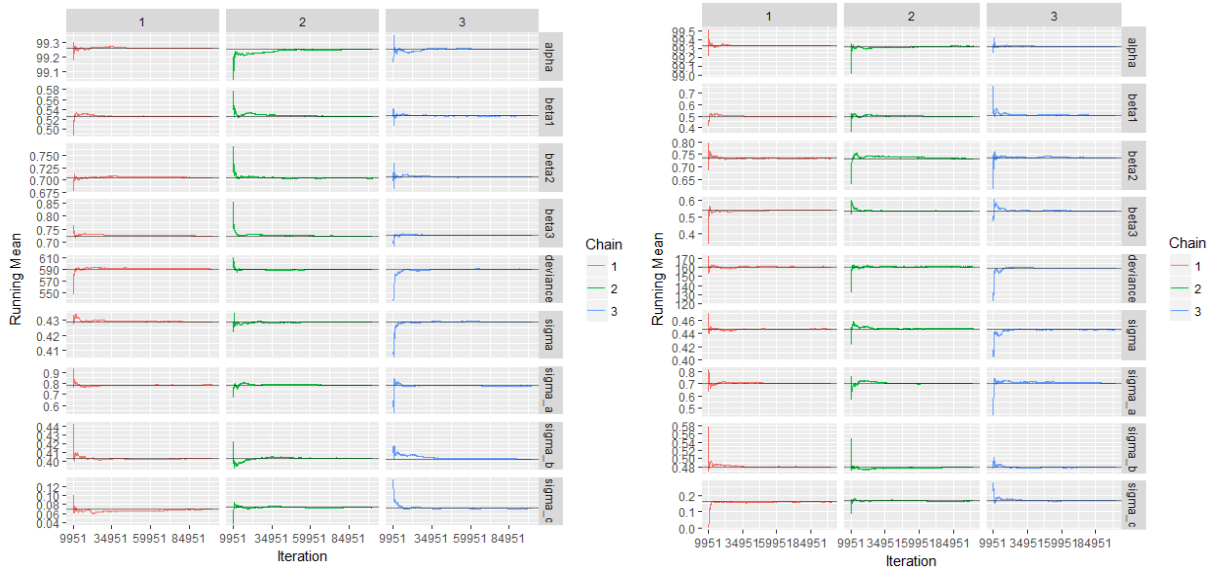
a convergence diagnostic check, it revealed that the initial values were easily forgotten in estimating the posterior. Running-mean plots for the various parameters of interest are shown in Figure 5 and indicate that stability was obtained almost immediately for all the parameters in all the three chains and in both datasets.



(a) Autocorrelation plot-complete data

(b) Autocorrelation plot-incomplete data

Figure 4: Autocorrelation plots



(a) Running-mean plot-complete data

(b) Running-mean plot-incomplete data

Figure 5: Running-mean plot

Formal convergence diagnostics included Geweke Z-test shown in Table 7 for each of the three chains. Approximately all the Z-values are within $[-1.96, 1.96]$, except for a few,

implying that the burn-in and the total chain were sufficient for both datasets.

Table 7: Geweke Z-values

| | Complete data | | | Incomplete data | | |
|-----------------------|---------------|---------|---------|-----------------|---------|---------|
| | Chain 1 | Chain 2 | Chain 3 | Chain 1 | Chain 2 | Chain 3 |
| α | -0.39 | -2.22 | -1.01 | 0.48 | -1.94 | 0.07 |
| β_1 | 2.10 | 1.96 | 0.51 | 1.59 | 0.19 | 1.62 |
| β_2 | 0.42 | 0.95 | 1.11 | 0.26 | 1.02 | 1.52 |
| β_3 | 2.45 | 1.83 | 1.91 | -0.56 | -0.096 | 1.28 |
| deviance | 1.65 | 0.12 | -0.31 | -0.42 | -0.48 | -0.02 |
| σ_{res} | 1.03 | -0.65 | 0.23 | -0.37 | 0.79 | -0.19 |
| σ_{batch} | -1.29 | 1.44 | -0.22 | 0.11 | 1.36 | 2.21 |
| $\sigma_{bag(batch)}$ | -0.24 | -0.86 | 2.16 | 1.36 | -1.83 | 0.01 |
| $\sigma_{A.run}$ | -1.19 | -0.19 | -0.06 | 0.12 | -0.33 | 0.42 |

The efficiency of the MCMC was also evaluated looking at the Monte-Carlo (MC) error which should be less than 5% of the respective standard deviations for good accuracy of the estimates. The MC error is shown in the Table 8 below. The MC error is lower than 5% of the standard deviation for all the parameters.

Table 8: MC error estimated via a time-series approach

| | Incomplete data | | | Complete data | | |
|-----------------------|-----------------|----------|--------------|---------------|----------|--------------|
| | SD | MC error | 5% of the SD | SD | MC error | 5% of the SD |
| α | 0.21 | 0.003 | 0.01 | 0.20 | 0.002 | 0.01 |
| β_1 | 0.15 | 0.002 | 0.007 | 0.04 | 0.001 | 0.002 |
| β_2 | 0.11 | 0.002 | 0.006 | 0.05 | 0.0001 | 0.002 |
| β_3 | 0.12 | 0.002 | 0.006 | 0.046 | 0.001 | 0.002 |
| deviance | 22.36 | 0.33 | 1.12 | 25.84 | 0.42 | 1.29 |
| σ_{res} | 0.04 | 0.001 | 0.002 | 0.014 | 0.0002 | 0.001 |
| σ_{batch} | 0.19 | 0.003 | 0.01 | 0.16 | 0.002 | 0.008 |
| $\sigma_{bag(batch)}$ | 0.05 | 0.001 | 0.003 | 0.028 | 0.0004 | 0.00140 |
| $\sigma_{A.run}$ | 0.10 | 0.002 | 0.01 | 0.04 | 0.001 | 0.002 |

4.2.3 Posterior predictive checks

Posterior predictive checks were used to assess normality of the random intercepts and of residuals. Discrepancy measures based on kurtosis and skewness were then obtained. In a hierarchical model, assumptions are made at each of the levels of hierarchy and posterior predictive checks based on Kolmogorov-Smirnov and Sinharay and Stern discrepancy measures were also applied. In checking normality of residuals, a PPP-value of 0.76 was obtained which is not small, say ≥ 0.05 . The PPP-values for the above discrepancy measures were obtained as follows for the three random intercepts:

Table 9: Assessing normality of random effects

| Random effect | Complete data | | | | Incomplete data | | | |
|----------------|---------------|--------|------|------|-----------------|--------|------|------|
| | D_skew | D_kurt | D_KS | D_SS | D_skew | D_kurt | D_KS | D_SS |
| batch | 0.66 | 0.46 | 0.49 | 0.85 | 0.48 | 0.54 | 0.52 | 0.48 |
| bag(batch) | 0.23 | 0.32 | 0.35 | 0.32 | 0.52 | 0.44 | 0.45 | 0.53 |
| analytical run | 0.51 | 0.51 | 0.50 | 0.52 | 0.47 | 0.52 | 0.47 | 0.44 |

None of the p-values is ≤ 0.05 , indicating normality of the random effects in the two datasets. Another model checking included check for outliers in the data which was done using posterior predictive ordinate (PPO). PPO plots for the two datasets are shown in Figure 6. A relatively small value for PPO may imply a possible outlying observation. Plots of inverse-CPO are also shown in Figure 8 (see Figure in appendix) with extreme values of the same flagged as possible outliers.

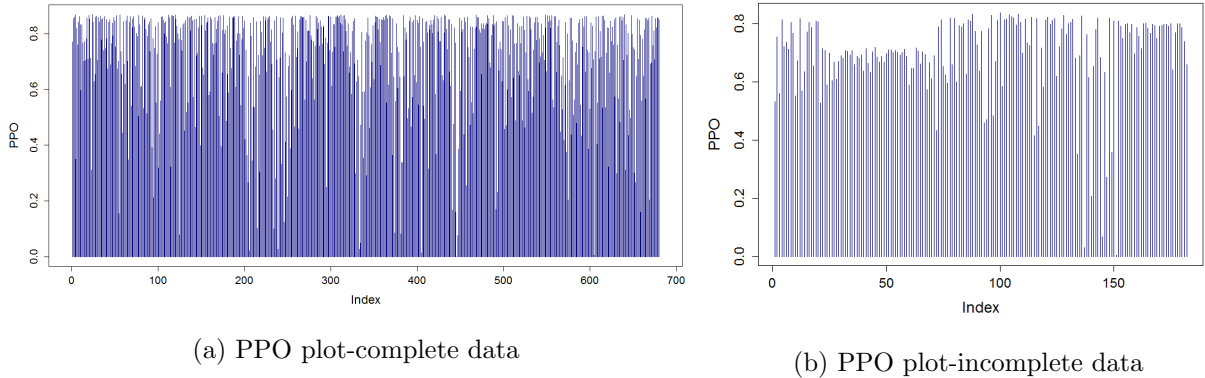


Figure 6: Posterior Predictive Ordinate plots

For the complete data, there was one outlying observation. A further look revealed that the observation was from batch 4 which, as was shown earlier, had the least assay on average compared to other batches. The assay for this observation was 96.75 which was indeed outlying for that batch. Nonetheless, it was still in acceptable range and was kept in the dataset. In the incomplete dataset, a few observations were outlying. These had assay below 98% but still above 95% and were also kept in the dataset. Recommendation, however, would be to ensure that there was perfect mixing in the process and correct recording of the assays obtained.

4.3 Variance Partition Coefficient

The total variation was split into several components. The table below shows the posterior median representing the proportion assigned to each of the components for the two datasets.

Table 10: Variance Partition Coefficient

| | Variance partition coefficient (VPC) | |
|----------------------|--------------------------------------|--------------------------|
| | Complete dataset | Incomplete dataset |
| | Posterior median[95% CI] | Posterior median[95% CI] |
| VPC (batch) | 0.46 [0.37, 0.57] | 0.39 [0.24, 0.53] |
| VPC (bag in batch) | 0.24 [0.19, 0.29] | 0.27 [0.20, 0.35] |
| VPC (analytical run) | 0.04 [0.004,0.08] | 0.09 [0.006,0.21] |
| VPC (residual) | 0.26 [0.20,0.30] | 0.25 [0.19, 0.32] |

From the results, the within-batch variability is lower than the between-batch variability. This is to be expected since assays in the same batch are produced much more together and have similar processing history. The assays within a batch are therefore more similar and the lower variation can be an indication of a good quality batch. The lowest source of variation was across analytical runs which indicates that even assays measured under different runs are not very different. Thus from the analytical run perspective, the process is good and not modeling this variation would not be very alarming.

4.4 Prediction of future batches

As aforementioned, there was interest to predict whether future batches of the product will pass a criteria imposed. This criteria was to declare a batch as acceptable if its mean assay was between 95% and 105%. The primary model was the R2jags model and prediction was done based on this model. Prediction was done conditional on all the random effects as well on a marginal batch level. The first case continued as follows; given a process mean of 100% (the ideal value), interest was to predict that a new composite sample taken from an existing batch, existing bag and analytical run would be acceptable by taking into account the random effect for the batch, bag in batch and the random effect for analytical run in the computation. The second case proceeded as follows; given a value for the batch mean, prediction was done for a future batch with 3 bags and assay measured with an existing analytical run. Several values for batch mean were used; 90, 95, 100, to evaluate behavior of future batches when there is a shift in the process mean. The batch was then assessed for acceptability. This was repeated for each of the MCMC iterations and the proportion of times the criteria was passed calculated. The table below shows these proportions for the complete as well as the incomplete data set.

Table 11: Proportion that passed acceptance criteria

| | Complete data | Incomplete data |
|-----------------------------------|---------------|-----------------|
| Conditional on all random effects | 1 | 1 |
| Given batch mean =100 | 1 | 1 |
| Given batch mean =95 | 0.12 | 0.13 |
| Given batch mean =90 | 0 | 0 |

The results indicate that conditional on all the random effects, a new composite sample would always pass the acceptance criteria for the incomplete and complete data. For a future batch with mean 95, there was a probability of 12% for future batches to be acceptable in the complete data and 13% for the incomplete data. In addition, if the batch mean was 90, such that the assay was produced with low quality raw materials, future batches would always never be acceptable. From these results, future batches can therefore be predicted on whether or not they would pass the acceptance criteria.

4.5 Secondary results

4.5.1 Results from frequentist and Brms on complete data

The jags model discussed above was the primary model of interest. In addition, there was an interest to explore the Brms package as an alternative to Bayesian model implementation. The formulation of the Brms package, however, does not provide a way to define a user-specified residual variability for the composite sample. The model was therefore fit without taking the design of the residual variability into account making it not directly comparable to the R2jags model. Nonetheless, to compare with the Brms model, a frequentist equivalent was fit also ignoring the design of the residual variability. It should be emphasized that while they both do not consider how the composite samples were obtained, they were used here for illustrative purposes to provide guidance on which is a better modeling approach and to yield more insight into the Brms package. Only the complete data was used in this comparison since Brms got stuck sampling for the incomplete data probably due to the incomplete design.

The results from these two approaches are shown in Table 12. The frequentist model was fit using maximum likelihood method. From this approach, the total variability was $0.38+0.38+0.72+0.09=1.57$, with the higher percentage (45.95%) attributed to the variance between batches. Effect of location on the assay was evaluated using a likelihood ratio test comparing a model with and without location as a covariate. A p-value < 0.0001 was obtained indicating that indeed location has an influence on the assay obtained with the lowest being for a sample taken at the beginning of the bag.

The results using Brms package are also shown in Table 12. Half-Cauchy (0,5) was used as an non-informative prior for the variance parameters in the model while for the "fixed effects" parameters, $N(0, 100^2)$ was used as a non-informative prior. The Half-Cauchy prior belongs to the half-t family of prior distributions that has more flexibility than the inverse-gamma family [23]. Convergence was achieved with 10000 iterations of three chains with a warm-up of 1000. Trace plots for the various parameters of interest are shown in Figure 8 (see Figure in Appendix). The full table of results from Brms is shown in Table 13 (see Table in Appendix).

Table 12: Frequentist and Brms for the complete data

| | Frequentist approach | Brms approach |
|--------------------------|----------------------|----------------------|
| | Estimate (s.e) | Estimate [95% CI] |
| α | 99.26 (0.18) | 99.26 [98.86, 99.66] |
| β_1 | 0.52 (0.05) | 0.52 [0.43, 0.62] |
| β_2 | 0.72 (0.04) | 0.72 [0.64, 0.80] |
| β_3 | 0.70 (0.04) | 0.70 [0.62, 0.79] |
| Sources of variability | | |
| σ_{Batch} | 0.70 | 0.78 [0.54, 1.17] |
| σ_{bag} | 0.38 | 0.38 [0.33, 0.44] |
| $\sigma_{Analyticalrun}$ | 0.09 | 0.08 [0.01, 0.17] |
| σ_{Res} | 0.38 | 0.38 [0.36, 0.41] |

The results from the two approaches are very similar. In both models, there was a significant effect of location with the beginning sample having the lowest percentage of API. Similarly, in both models, higher percentage of variation was attributed to the variance between batches while the lowest was between analytical runs.

5 Discussion

The evaluation of the sources of variability helps in evaluating the quality of the manufacturing process which ensures that end-users of the products are protected. In this case, variability was due to batches, bags within batches, analytical run as well as measurement error. Two datasets were considered to assess the parameters change when there is a complete and incomplete design. In practice, however, the incomplete design dataset reflects reality. The response was the percentage of active pharmaceutical ingredient in a sample (assay). Assay obtained as composite was an equal mixture of assay obtained from the beginning, middle and end indicating inherently different variability for the composite assays. The model was implemented in R using packages; R2jags, brms and lme4. Flexibility of the jags model allowed taking the actual structure of the data into account in a less complex way.

The final model chosen was the jags model fitted to the incomplete data. The flexibility of this model allowed easier definition of the different variability for samples taken from composite location. Results of the jags model indicated an effect of location on the assay since the 95% CI did not include zero. There was also larger between-batch variability than within-batch. The smaller within-batch variability is a good indication that the product is more similar within batches. The larger between batch variability implies that the batches are not consistent and they can be very different from each other. Halting of production between batches combined with machine restructuring contribute to the high variation between batches. This variability has to be taken into account especially in predicting whether future batches will be acceptable. The high variability across batches can cause huge inconsistencies in the efficacy of drugs that can have serious negative impacts on end users. The results are consistent with [11] who showed that variation across batches is usually the most prevalent.

To reduce this, pharmaceutical companies should try and make the production of batches to be as similar as possible, in terms of the machines involved etc. While there did not show to be any definite time trend across the batches, it may be important to produce the batches within short periods of time from each other and to ensure reconfiguration of the machines between batches is done with care. Another option would be to switch from batch production to continuous production. This can reduce quality variations as well as save on resources. The continuous batch production is, however, more complex in drug powder production due to the various steps required and remains less used in this industry. With the batch production method still highly used, it is then important to take the variability across batches into account in modeling. Failure to do so could lead to gross underestimation of process variation. The least contributor to the total variation was between analytical runs. This is a good indication since the analytical run variation is not very alarming. In general, consistency in every step of the manufacturing process should be upheld to reduce variations in the end product.

Further interest in the pharmaceutical industry is to ensure that future batches meet accep-

tance criteria set. The ideal assay is 100% but the batch is acceptable if its mean is between 95% and 105%. Prediction can therefore be done conditional on all random effects and also marginal on the batch level for future batches. It was revealed that a new composite sample, taken from an existing batch and measured by an existing analytical run, would always be acceptable for both complete and incomplete data. On the other hand, a future batch with mean equal to 90 such that it was produced with low quality raw materials and with composite samples measured by an existing analytical run, would always never be acceptable.

Secondary results were based on gaining more insight into the Brms package. The Brms got stuck fitting to the incomplete data which could be as a result of the incomplete design. Thus, only the complete dataset was used in this case. From the results there was a significant effect of location which was also concluded from the results of the frequentist evaluated using a likelihood ratio test with p-value $< .0001$. From both models, variation between batches explained much of the total variation. The between batch variability obtained under the Brms was, however, slightly higher than that obtained from frequentist. Nonetheless, with a balanced design and a non-complex model, the Brms is more advantageous as compared to the frequentist. Brms uses a Bayesian approach which offers better interpretability in terms of probability and represents parameter estimates with a distribution rather than just point estimates as in the frequentist. Brms is based on Hamiltonian Monte Carlo (HMC) which as aforementioned converges faster than MCMC. This was ascertained by the fact that it took only 10,000 iterations to achieve stability for the parameter estimates.

A Bayesian approach allows taking into account all the uncertainty about the model parameters. Each parameter, therefore, has a distribution defined by the prior distribution that captures uncertainty about that parameter [24]. Conventional methods may fail completely when it comes to complex models and the flexibility of Bayesian approach becomes even more appealing. As shown, the jags model formulation offers more flexibility as compared to Brms and frequentist approaches. In terms of the complete and incomplete design of the datasets, jags family has more flexibility as compared to Brms since as shown, Brms did not work for the incomplete design. There was an aim to explore the parameter changes when dealing with an incomplete and complete design. With respect to this, same qualitative results were obtained with slight quantitative changes in the parameter values. In terms of prediction, similar conclusions were made using both data sets.

In general, location was found to have an effect on the percentage of API. This means that it actually does matter where in the bag the assay was obtained. Based on the estimates for the parameters representing location, higher assay was obtained from the middle location as compared to the beginning of the bag. Of all the locations, sample obtained at the beginning had the least assay. This was also ascertained by the exploratory analysis. Perfect mixing of the powdered ingredients should be enhanced to ensure that there is uniformity in the contents of the powder.

6 Limitation

The model fitted under Brms package did not work for an unbalanced design dataset. This is unfortunate since most of the datasets in practice tend to have such designs. Further, the model formulation in Brms is somewhat rigid and it is not clear how to include a user-specified residual variance for different groups. From a practical point of view, Bayesian implementations take more computational time but due to advancement in efficient computer processors, this limitation is now almost not an issue.



7 Conclusion and recommendation

This study provided insights into the importance of modeling variation for quality manufacturing as well as on Bayesian approaches in different packages. Bayesian modeling has revolutionized how we approach complex models. As a recommendation, the pharmaceutical industry should consider applying such approaches as they could have huge positive impact on research. In addition, with a balanced or unbalanced design, jags model offers more flexibility compared to other models considered here. With a non-complex model and a balanced design, Brms offers an alternative to Bayesian modeling especially because it is not programming-intensive. Future developments should, however, seek to explore why Brms does not work for an unbalanced design.



References

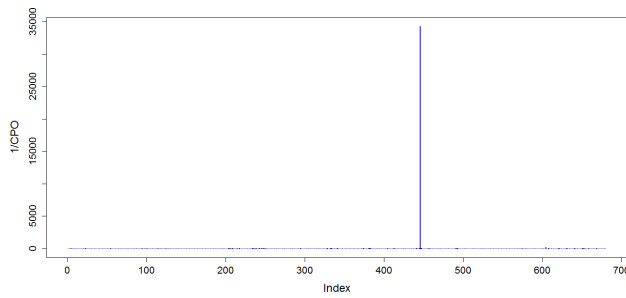
- [1] Rantanen, J., & Khinast, J. (2015). The future of pharmaceutical manufacturing sciences. *Journal of pharmaceutical sciences*, *104*(11), 3612-3638.
- [2] Tait, K. (1998). Pharmaceutical industry. *Encyclopaedia of Occupational Health and Safety*, 79-4.
- [3] Bumpas, J., & Betsch, E. (2009). Exploratory study on active pharmaceutical ingredient manufacturing for essential medicines. *Health, nutrition and population discussion paper. Washington, DC, International Bank for Reconstruction and Development and World Bank*.
- [4] Aundhia, C. J., Raval, J. A., Patel, M. M., Shah, N. V., Chauhan, S. P., Sailor, G. U., ... & Mahashwari, R. A. (2011). Spray drying in the pharmaceutical industry—a review. *Indo American Journal of Pharmaceutical Research*, *2*(1), 125-138.
- [5] Patil, J. (2016). Spray-Drying: An Emerging Technique for Pharmaceutical Product Development. *Journal of Pharmacovigilance*.
- [6] Khadka, P., Ro, J., Kim, H., Kim, I., Kim, J. T., Kim, H., ... & Lee, J. (2014). Pharmaceutical particle technologies: An approach to improve drug solubility, dissolution and bioavailability. *asian journal of pharmaceutical sciences*, *9*(6), 304-316.
- [7] Meeus, J., Lenaerts, M., Scurr, D. J., Amssoms, K., Davies, M. C., & Roberts, C. J. (2015). The influence of spray-drying parameters on phase behavior, drug distribution, and in vitro release of injectable microspheres for sustained release. *Journal of pharmaceutical sciences*, *104*(4), 1451-1460.
- [8] Spilker, B. (1994). *Multinational pharmaceutical companies: principles and practices*. Lippincott Williams & Wilkins, 2nd edition.
- [9] Kou, D., Wigman, L., Yehl, P., & Dong, M. W. (2015). Separation Science in Drug Development, Part IV: Quality Control. *LC GC NORTH AMERICA*, *33*(12), 900-+.
- [10] Hugo, J. (2012). *Bayesian tolerance intervals for variance component models* (Doctoral dissertation, University of the Free State).
- [11] Mockus, L., Peterson, J. J., Lainez, J. M., & Reklaitis, G. V. (2014). Batch-to-batch variation: a key component for modeling chemical manufacturing processes. *Organic Process Research & Development*, *19*(8), 908-914.
- [12] Kruschke, J. K., & Vanpaemel, W. (2015). Bayesian estimation in hierarchical models. *The Oxford handbook of computational and mathematical psychology*, 279-299.
- [13] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.
- [14] Lesaffre, E., & Lawson, A. B. (2012). *Bayesian biostatistics*. John Wiley & Sons.

-
- [15] Ravenzwaaij, D., Cassey, P., & Brown, S. D. (2018). A simple introduction to Markov Chain Monte–Carlo sampling. *Psychonomic bulletin & review*, 25(1), 143-154.
- [16] Neal, R. M. (2011). MCMC using Hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo*, 2(11).
- [17] Monnahan, C. C., Thorson, J. T., & Branch, T. A. (2017). Faster estimation of Bayesian models in ecology using Hamiltonian Monte Carlo. *Methods in Ecology and Evolution*, 8(3), 339-348.
- [18] Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.
- [19] Kaplan, D., & Depaoli, S. (2013). Bayesian statistical methods. *Oxford handbook of quantitative methods*, 407-437.
- [20] Ntzoufras, I. (2009). *Bayesian modeling using WinBUGS*. John Wiley & Sons.
- [21] Sinharay, S., & Stern, H. S. (2003). Posterior predictive model checking in hierarchical models. *Journal of Statistical Planning and Inference*, 111(1-2), 209-221.
- [22] Spiegelhalter, D. J., Abrams, K. R., & Myles, J. P. (2004). *Bayesian approaches to clinical trials and health-care evaluation* (Vol. 13). John Wiley & Sons.
- [23] Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian analysis*, 1(3), 515-534.
- [24] van de Schoot, R., & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. *European Health Psychologist*, 16(2), 75-84.

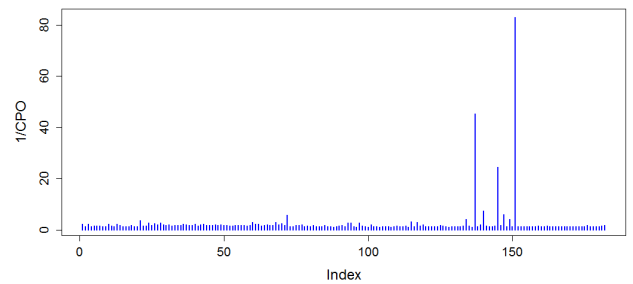
Appendix

Table 13: Parameter estimates for BRMS for the complete dataset

| | Estimate | Est.Error | l-95% CI | u-95% CI | Eff.Sample | Rhat |
|-----------------------|----------|-----------|----------|----------|------------|------|
| Intercept | 99.26 | 0.20 | 98.86 | 99.66 | 2255 | 1.00 |
| SampleComposite | 0.52 | 0.05 | 0.43 | 0.62 | 12853 | 1.00 |
| SampleEnd | 0.72 | 0.04 | 0.64 | 0.80 | 17061 | 1.00 |
| SampleMiddle | 0.70 | 0.04 | 0.62 | 0.79 | 15512 | 1.00 |
| σ_{res} | 0.38 | 0.013 | 0.36 | 0.41 | 12790 | 1.00 |
| σ_{Batch} | 0.78 | 0.16 | 0.54 | 1.17 | 5239 | 1.00 |
| $\sigma_{bag(batch)}$ | 0.38 | 0.028 | 0.33 | 0.44 | 7436 | 1.00 |
| $\sigma_{A.run}$ | 0.09 | 0.04 | 0.01 | 0.17 | 3035 | 1.00 |



(a) Inverse-CPO plot-complete data



(b) Inverse-CPO plot-incomplete data

Figure 7: Index plot of 1/CPO

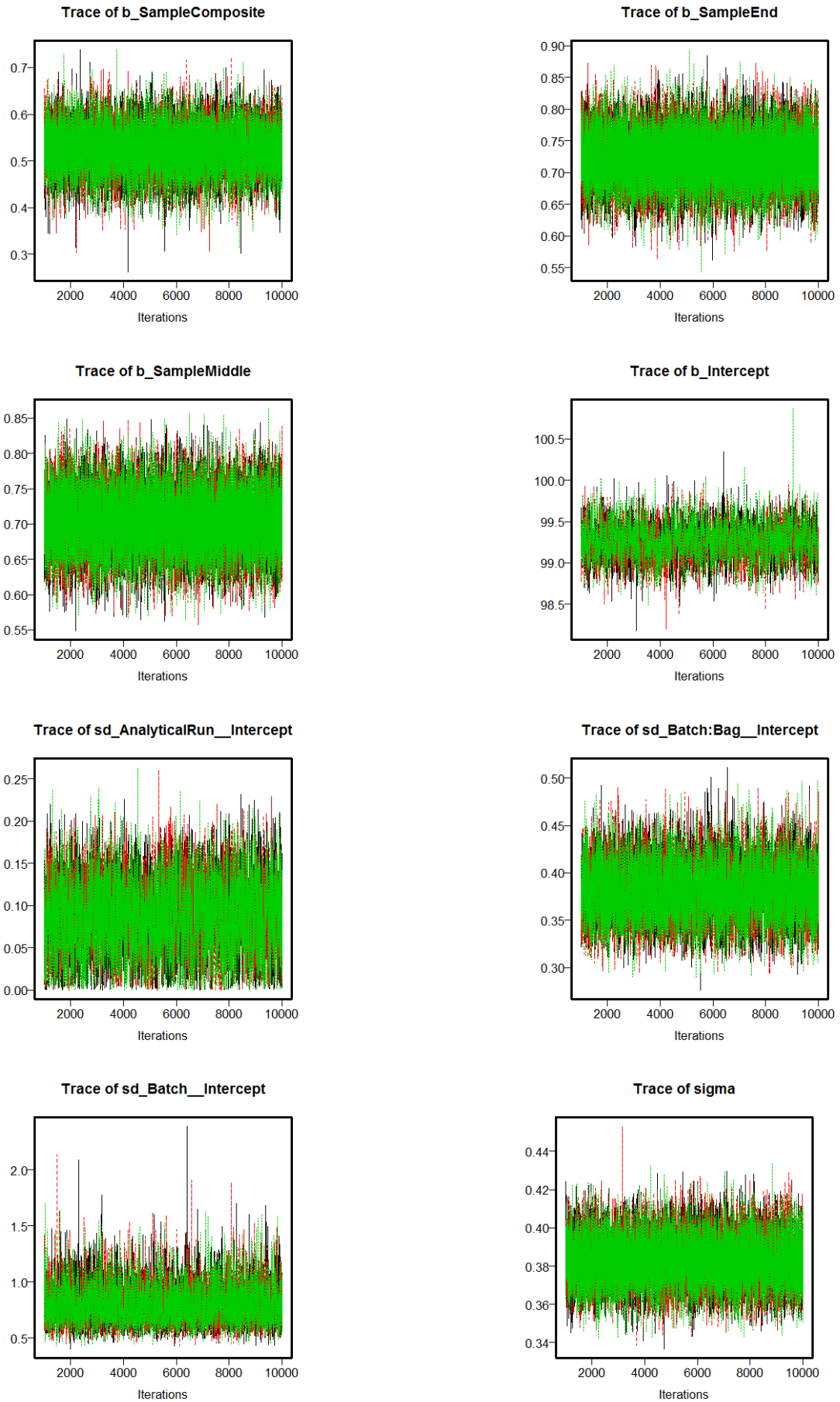


Figure 8: Trace plot for the BRMS parameters

R codes

```
#####
# Reading data #
#####
# complete data
powder=read.csv("C:\\Users\\VIVIENN\\Desktop\\HASSELT UNIVERSITY\\
2nd year\\thesis
\\dataSetFull.csv",header=TRUE)
# incomplete data
powder_incomp=read.csv("C:\\Users\\VIVIENN\\Desktop\\HASSELT UNIVERSITY\\
2nd year
\\thesis\\dataSet.csv",header=TRUE)

#####
# R2jags for the complete data-for the incomplete data, its just changing the data
#
#####
powder$Composite1=ifelse(Sample=="Composite",1,0)
powder$Bag1=Batch:Bag
powder$Bag1=as.numeric(as.factor(powder$Bag1))
powder$AnalyticalRun=as.numeric(as.factor(powder$AnalyticalRun))
powder$Batch=as.numeric(as.factor(powder$Batch))

model_jags=function(){
for (i in 1:N){
Assay[i]~dnorm(mu[i],tau1[i])
mu[i]=alpha+ beta1*Sample1[i]+beta2*Sample2[i]+beta3*Sample3[i]+a[Batch[i]]+
b[Bag1[i]]
+c[AnalyticalRun[i]]
Assay.rep[i]~dnorm(mu[i],tau1[i])
tau1[i]=(3*tau*Composite1[i] + tau*(1-Composite1[i]))

#ordinary residual
r[i]=Assay[i]-mu[i]
r.rep[i]=Assay.rep[i]-mu[i]
# standardized residual
rs[i]=r[i]/(1/sqrt(tau))
rs.rep[i]=r.rep[i]/(1/sqrt(tau))
# checking skewness and kurtosis of the residuals
m3[i]=pow(rs[i],3)
m4[i]=pow(rs[i],4)
m3.rep[i]=pow(rs.rep[i],3)
```

```

m4.rep[i]=pow(rs.rep[i],4)
}
m3s=sum(m3[])/N
m4s=sum(m4[])/N-3
m3s.rep=sum(m3.rep[])/N
m4s.rep=sum(m4.rep[])/N-3
skewness.test=step(m3s.rep-m3s)
kurtosis.test=step(m4s.rep-m4s)

#priors
alpha~dnorm(0,0.0001)
for (j in 1:Nbatch){
a[j]~dnorm(0,tau_a)
a.rep[j]~dnorm(0.0,tau_a)
m3.a[j]= pow((a[j])/sigma_a,3)
m4.a[j]=pow((a[j])/sigma_a,4)
m3.a.rep[j]= pow((a.rep[j])/sigma_a,3)
m4.a.rep[j]=pow((a.rep[j])/sigma_a,4)
}
#normality of random batches effect
tmin = min(a)
tmax=max(a)
tmin.rep= min(a.rep)
tmax.rep =max(a.rep)
tmin.test = step(tmin.rep-tmin)
tmax.test= step(tmax.rep-tmax)

m3a=sum(m3.a[])/Nbatch
m4a=sum(m4.a[])/Nbatch-3
m3a.rep=sum(m3.a.rep[])/Nbatch
m4a.rep= sum(m4.a.rep[])/Nbatch-3
skewness.a.test=step(m3a.rep - m3a)
kurtosis.a.test=step(m4a.rep-m4a)
#ksmethod
sorta <- sort(a[])
sorta_new <- sort(a.rep[])
  for (j in 1:Nbatch) {
    pa[j] <- phi(sorta[j])
    panew[j] <- phi(sorta_new[j])
    maxpa[j] <- max(pa[j] - ((j-1)/Nbatch) , (j/Nbatch) - pa[j])
    maxpanew[j] <- max(panew[j] - ((j-1)/Nbatch) , (j/Nbatch) - panew[j])
  }
  ks_a <- max(maxpa)

```

```

        ks_aneu <- max(maxpaneu)
        test5 <- step(ks_aneu - ks_a)
#sinharay and stern
        absa <- abs(tmax - sorta[50]) - abs(tmin - sorta[50]) #50 the median
        absa_new <- abs(tmax.rep - sorta_new[50]) - abs(tmin.rep - sorta_new[50])
        test6 <- step(absa_new - absa)

# Between batches variance
        sigma_a~ dunif (0, 100)
var_a<- pow(sigma_a,2)
        tau_a<- pow(sigma_a, -2)

for (k in 1:(Nbags)){
b[k]~dnorm(0,tau_b)
b.rep[k]~dnorm(0,tau_b)
m3.b[k]= pow((b[k])/sigma_b,3)
m4.b[k]=pow((b[k])/sigma_b,4)
m3.b.rep[k]= pow((b.rep[k])/sigma_b,3)
m4.b.rep[k]=pow((b.rep[k])/sigma_b,4)

}
#normality of random bags effect
tmin1 = min(b)
tmax1=max(b)
tmin.rep1= min(b.rep)
tmax.rep1 =max(b.rep)
tmin.test1 = step(tmin.rep1-tmin1)
tmax.test1= step(tmax.rep1-tmax1)

m3b=sum(m3.b[])/(Nbags)
m4b=sum(m4.b[])/(Nbags)-3
m3b.rep=sum(m3.b.rep[])/(Nbags)
m4b.rep= sum(m4.b.rep[])/(Nbags)-3
skewness.b.test=step(m3b.rep-m3b)
kurtosis.b.test=step(m4b.rep-m4b)
#ksmethod
        sortb <- sort(b[])
        sortb_new <- sort(b.rep[])
        for (k in 1:(Nbags)) {
                pb[k] <- phi(sortb[k])
                pbnew[k] <- phi(sortb_new[k])
maxpb[k] <- max(pb[k] - ((k-1)/(Nbags)) , (k/(Nbags)) - pb[k])
maxpbnew[k] <- max(pbnew[k] - ((k-1)/(Nbags)) , (k/(Nbags)) - pbnew[k])

```

```

    }
    ks_b <- max(maxpb)
    ks_bnew <- max(maxpbnew)
    test7 <- step(ks_bnew - ks_b)
#sinharay and stern
    absb <- abs(tmax1 - sortb[50]) - abs(tmin1 - sortb[50]) #50 the median
    absb_new <- abs(tmax.rep1 - sortb_new[50]) - abs(tmin.rep1 - sortb_new[50])
    test8 <- step(absb_new - absb)

# Between bags variance
    sigma_b ~ dunif(0, 100)
var_b <- pow(sigma_b, 2)
    tau_b <- pow(sigma_b, -2)
for (l in 1:NRun){
c[l] ~ dnorm(0, tau_c)
c.rep[l] ~ dnorm(0.0, tau_c)
m3.c[l] = pow((c[l])/sigma_c, 3)
m4.c[l] = pow((c[l])/sigma_c, 4)
m3.c.rep[l] = pow((c.rep[l])/sigma_c, 3)
m4.c.rep[l] = pow((c.rep[l])/sigma_c, 4)
}
#normality of random runs effect
tmin2 = min(c)
tmax2 = max(c)
tmin.rep2 = min(c.rep)
tmax.rep2 = max(c.rep)
tmin.test2 = step(tmin.rep2 - tmin2)
tmax.test2 = step(tmax.rep2 - tmax2)

m3c = sum(m3.c[])/NRun
m4c = sum(m4.c[])/NRun-3
m3c.rep = sum(m3.c.rep[])/NRun
m4c.rep = sum(m4.c.rep[])/NRun-3
skewness.c.test = step(m3c.rep - m3c)
kurtosis.c.test = step(m4c.rep - m4c)
#ksmethod
    sortc <- sort(c[])
    sortc_new <- sort(c.rep[])
    for (l in 1:NRun) {
pc[l] <- phi(sortc[l])
pcnew[l] <- phi(sortc_new[l])
maxpc[l] <- max(pc[l] - ((1-l)/NRun), (1/NRun) - pc[l])
maxpcnew[l] <- max(pcnew[l] - ((1-l)/NRun), (1/NRun) - pcnew[l])

```

```

    }
    ks_c <- max(maxpc)
    ks_cnew <- max(maxpcnew)
    test9 <- step(ks_cnew - ks_c)
#sinharay and stern
  absc <- abs(tmax2 - sortc[50]) - abs(tmin2 - sortc[50]) #50 the median
  absc_new <- abs(tmax.rep2 - sortc_new[50]) - abs(tmin.rep2 -
    sortc_new[50])
  test10 <- step(absc_new - absc)

# Between analytical run variance
  sigma_c~ dunif (0, 100)
var_c<- pow(sigma_c,2)
  tau_c<- pow(sigma_c, -2)

beta1~dnorm(0,0.0001)
beta2~dnorm(0,0.0001)
beta3~dnorm(0,0.0001)
sigma~dunif(0,100)
tau=1/(sigma*sigma)
var1=1/(tau)
#####CALCULATION OF PPO
for (i in 1:N){
log.ppo[i] <- -0.5 * log( 2*3.14 ) + 0.5*log(tau)-0.5*tau*(Assay[i]-mu[i])*
(Assay[i]-mu[i])
ppo[i]<- exp(log.ppo[i])
icpo[i]<- 1/ppo[i]
}
#tests
test[1]=skewness.test
test[2]=kurtosis.test
test[3]=skewness.b.test
test[4]=kurtosis.b.test
test[5]=tmin.test
test[6]=tmax.test
test[7]=skewness.a.test
test[8]=kurtosis.a.test
test[9]=skewness.c.test
test[10]=kurtosis.c.test
test[11]=tmin.test1
test[12]=tmax.test1
test[13]=tmin.test2
test[14]=tmax.test2

```

```

test[15]=test5
test[16]=test6
test[17]=test7
test[18]=test8
test[19]=test9
test[20]=test10
#Variance partition
vp=sigma/(sigma_a+sigma_b+sigma_c+sigma)
vpc=sigma_a/(sigma_a+sigma_b+sigma_c+sigma)
vpc1=sigma_b/(sigma_a+sigma_b+sigma_c+sigma)
vpc2=sigma_c/(sigma_a+sigma_b+sigma_c+sigma)
}
###initial values#
init_values=function(){
list(alpha=rnorm(1),sigma_a=runif(1),beta1=rnorm(1),beta2=rnorm(1),
beta3=rnorm(1),sigma=runif(1),sigma_b=runif(1),sigma_c=runif(1))
}
params= c("alpha", "beta1","beta2","beta3","sigma","sigma_a","sigma_b",
"sigma_c","test","vp","vpc","vpc1","vpc2")
library(R2jags)
set.seed(123)
#The implementation
jagsdata_s31st <- with(powder, list(Assay=Assay,Sample1=Sample1,
Sample2=Sample2,Sample3=Sample3, Batch = Batch ,Bag1=Bag1,
AnalyticalRun=AnalyticalRun,
  N=N,Composite1=Composite1, Nbatch = Nbatch, Nbags=Nbags, NRun=NRun))
finalmodel <- jags(data =jagsdata_s31st, inits = init_values,
parameters.to.save = params, model.file = model_jags1st,n.chains = 3, n.thin=50,
n.iter = 100000, n.burnin = 10000)

finalmodel$BUGSoutput$summary
#####
# Model diagnostics #
#####
library(coda)
finalmodel_fit=as.mcmc(finalmodel)
summary(finalmodel_fit)
library(ggmcmc)
myjags_ggs <- ggs(finalmodel_fit)
ggs_autocorrelation(myjags_ggs)
ggs_traceplot(myjags_ggs)
ggs_running(myjags_ggs)
ggs_geweke.plot(myjags_ggs)

```

```

#####
# prediction #
#####
jagsmatrix=as.matrix(finalmodel$BUGSoutput$sims.matrix)
bet1=data.frame(jagsmatrix[,2]) #dataframe containing beta corresponding
#to composite
bet2=data.frame(jagsmatrix[,6]) #dataframe containing sigma
bet3=data.frame(jagsmatrix[,7] )#dataframe containing sigma_batch
bet4=data.frame(jagsmatrix[,8] )#dataframe containing sigma_bag
bet5=data.frame(jagsmatrix[,9]) #dataframe containing sigma_analyticalrun
y3=matrix(ncol=3,nrow=5400)
for(i in 1:5400){
y3[i,]<-100+bet1[i,]+rnorm(1,0,bet3[i,])+rnorm(3,0,bet2[i,])+
rnorm(3,0,bet4[i,])+
rnorm(1,0,bet5[i,])
}
pass3=ifelse((y3[,1]>=95 & y3[,2]<=105),"yes","no")
pass4=as.data.frame(pass3)
pass4[pass4[,c(1,2,3)]=="yes",]
library(dplyr)
filter(pass4,V1=="yes",V2=="yes",V3=="yes")
5400/5400 # this is 95-105
#####
### for a specified batch mean #
#####
y4=matrix(ncol=3,nrow=5400)
for(i in 1:5400){
y4[i,]<-90+rnorm(3,0,bet2[i,])+rnorm(3,0,bet4[i,])+
rnorm(1,0,bet5[i,])
}
pass5=ifelse((y4[,1]>=95 & y4[,2]<=105),"yes","no")
pass6=as.data.frame(pass5)
library(dplyr)
filter(pass6,V1=="yes",V2=="yes",V3=="yes")
#####
### BRMS #
#####
library(brms)
prior<-c(set_prior("normal(0,100)", class="b"),#prior for the beta's
set_prior("cauchy(0,5)", class="sigma"),#prior for the residual std. deviation
set_prior("cauchy(0,5)", class="sd"))
model_stan=brm(Assay~Sample+(1|Batch/Bag)+(1|AnalyticalRun)

```

```
,prior=prior,data=powder,family=gaussian(),warmup=1000,iter=10000,chains=3)
print(model_stan,digits=4)
summary(model_stan)
autocorr.plot(stan_model)
traceplot(stan_model)
#####
# Frequentist #
#####
library(lme4)
fm11=lmer(Assay~Sample+(1|Batch1/Bag)+(1|AnalyticalRun),powder, REML=FALSE)
summary(fm11)
```


Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
Bayesian hierarchical models for powder manufacturing

Richting: **Master of Statistics-Biostatistics**

Jaar: **2018**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Weru, Vivienn

Datum: **15/06/2018**