2014•2015
FACULTY OF SCIENCES
*Master of Statistics*

# Master's thesis
## Statistical models for stability studies


Supervisor :
Prof. dr. Roel BRAEKERS

Supervisor :
Dr. LUWIS DIYA

Germaine Uwimpuhwe


*Thesis presented in fulfillment of the requirements for the degree of Master of Statistics*


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

universiteit hasselt
KNOWLEDGE IN ACTION

Universiteit Hasselt | Campus Hasselt | Martelarenlaan 42 | BE-3500 Hasselt
Universiteit Hasselt | Campus Diepenbeek | Agoralaan Gebouw D | BE-3590 Diepenbeek

universiteit hasselt | Maastricht University

2014•2015
FACULTY OF SCIENCES
*Master of Statistics*

# Master's thesis

## Statistical models for stability studies

Supervisor :
Prof. dr. Roel BRAEKERS

Supervisor :
Dr. LUWIS DIYA

Germaine Uwimpuhwe
*Thesis presented in fulfillment of the requirements for the degree of Master of Statistics*

universiteit hasselt | Maastricht University

**Abstract**

Stability studies are conducted at all phases of the drug development cycle, with the main objective of having a stable product on market. In this project we aimed at evaluating if the shelf life could be extended from 24 (current shelf life) to 36 months, quantifying pharmaceutical stability such as shelf life, release limit, degradation rate ( annually and at the end of both shelf lives) and consumer/producer risk. The assay data are longitudinal from 50 different batches, which were put in stability chamber within two storage conditions and monitored up to 24 months. Linear mixed model in frequentist(classical and quantile regression) and bayesian approach were fitted. From exploratory data analysis and model reduction, random intercept model with linear mean structure was selected. The highest degradation rate after one, two and three years were from API B at storage condition 25C /60%RH. Given a shelf life of 24 (36) months, the release limit for API A were 96.94 (97.37) for batches mean , and 97.37 (97.37) for individual tablets. The similar analysis was done for API B. The estimated producer and consumer risk were around zero for a shelf life of 24 month and bit high (9%) for shelf life of 36 months at storage condition 30C/75%RH. In summary the shelf life for the drug stored at condition 25C/60%RH can be extended to 36 months which was not the case for condition 30C/75%RH.

# Contents

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my external and internal supervisors, Dr. Luwis Diay and Prof. Dr. Roel Braekers for their guidance, patience and support throughout this thesis. It is my privilege to thank all my family members, especially to my lovely husband Theophile for his advices and being the source of my motivation. I would also like to thank all my classmates for their constant help and encouragement, but more especially to my group members. I am also grateful to thank all my friends (Hiwot) for sharing their generous ideas and precious time. Above all, To the Almighty God, the author of knowledge and wisdom, for his limitless love.

Thank you.

# 1 Introduction

A drug product is comprises of the Active Pharmaceutical Ingredients (API) and some excipients. Several attributes of a drug product are assessed to establish the quality profile of a given drug, for instance stability of the dug substances/APIs within a drug product, dissolution profiles, color, pH, etc. These attributes are usually referred to as the Critical Quality Attributes (CQA). A major CQA is the stability of the APIs which constitute a drug product over time. In a stability study the different the drug product is put in stability chambers which represent different environmental factors such as temperature, humidity and light. These environmental factors usually represent different climatic zones were the drug will be distributed or marketed.

One major characteristic of a drug product is the stability profile of the active pharmaceutical ingredient (API) over time. The drug product is considered unsafe, if it fails to remain within the approved specifications for the identity, strength, quality, and purity. Pharmaceutical companies usually conduct stability studies to provide evidence on how the quality of a drug substance or drug product varies with time under the influence of a variety of environmental factors (temperature, humidity, and light), to establish a re-test period for the drug substance or a shelf life for the drug product and recommended storage conditions(Shein-chung et al 2007).

According to the International Committee Of Harmonization (ICH), shelf life (expiration dating period) is "the time period during which a drug product is expected to remain within the approved shelf life specification, provided that it is stored under the conditions defined on the container label." So a stability study should establish "with a high degree of confidence, a re-test period or shelf life during which a quantitative attribute will remain within acceptance criteria for all future batches manufactured, packaged, and stored under similar circumstances (Q1E). According to the FDA(1987, 1998), the time at which the average drug characteristic remains within an approved specification after manufacture is recommended as the shelf-life of the drug. Shelf-life can be estimated based on assay results of the drug characteristics from a stability study conducted during the process of drug development with respect to the International Conference of Harmonization (ICH) Tripartite Guidelines for Q1E (Shein-chung et al 2007). Following current regulatory practice of the FDA, the shelf life can not exceed 12 months

beyond the last stability testing (follow-up time) when the 95 to 105% of label claim are used as the specification limit.

Given the shelf life and the specification limit, a lot released with results belonging to the release limit (or interval) will likely to remain within the specification limits at the end of shelf life. The idea is to know how much ingredients to put in drug product to guarantee quality at the end of shelf life. Producer or manufacture risk is the probability that a good product falsely determined as violating the approval specification limits at release time. The opposite is the consumer risk which is the probably of accepting a defective product as good.

Both the FDA and ICH require that at least three batches, and preferably more should be tested to allow account for batch-to-batch variability so that a single shelf-life is applicable to all future batches manufactured under similar circumstances. Usually fixed effects models by batch are fitted and the minimum shelf life is taken as the shelf life of the drug product (Shein-chung et al 2007). This approach has limitations in that the fixed effects models are used to make a statement about the given batches and there is no optimal usage of information. Mixed effects models allows for the pooling of information across batches to come up with the optimal shelf life estimates as it properly adjust for batch-to batch variability. Another benefit of the mixed effects model is that it generalizes to future batches, not necessarily limited to the current observed batches.

In order to determine the rates of chemical and physical reactions and their relationships with environmental factors on the API, accelerated, intermediate and long-term studies are usually conducted. Accelerated and intermediate studies are so-called short-term stability studies conducted under exaggerated conditions while long-term studies simulate the actual packaging and conditions used for storage (Hsin-ya et al, 2010). The 1987 FDA stability guideline suggests that stability testing be done every three months during the first year, six months intervals during the second year, and annually thereafter. However, if a drug product is expected to degrade rapidly, more frequent sampling are necessary.

The aim of this project is to calculate the shelf life of drug product X (due to confidentiality the name of the drug was not given), calculate the annual rate of change for the two API in drug X (A and B), calculate the release limits and the producer and consumer risk. The current shelf life of drug X is 24 months. There is an interest to evaluate if the shelf life can be extended to 36 months. So the main objective is to ascertain if the shelf lives of the two APIs exceed 36 months thus allowing for the shelf life to be established as 36 months.

This project aimed at assessing different statistical models in quantifying pharmaceutical stability, to calculate the shelf life of drug product X (due to confidentiality the name of the drug was not given), calculate the annual rate of change for the two API in drug X (A and B), calculate the release limits and the producer and consumer risk. The current shelf life of drug X is 24 months. There is an interest to evaluate if the shelf life can be extended to 36 months. So the main objective is to ascertain if the shelf lives of the two APIs exceed 36 months thus allowing for the shelf life to be established as 36 months.

This thesis is organized as follows: Section 2 gives a brief overview of the used methodology. In Section 3, we illustrate the application to the data at hand. We end with conclusions and discussion with some recommendation in Section 4

## 1.1 Data Description

The data considered in this thesis are from a longitudinal for two APIs ( A and B). At release (time=0 months), 50 registered lots were monitored. Only 30 lots were put on stability chamber at storage conditions 25C/60%RH and 30C/75%RH and monitored at stability time 1,3,6,12,18 and 24 months. In total, the data set consisted of 820 assay (% of the label claim). The two storage conditions correspond to the two different climatic zones. 25C/60%RH match with subtropical and mediterranean climate, while 30C/75%RH match with hot and very humid climate.

All analyses were performed using R (R Core Team (2015)) . All hypotheses were tested at 5% significance level for the frequentist approaches reported in this thesis.

# 2 Methodology

Linear mixed models were used to evaluate the stability profiles of the two APIs in drug X. In stability studies classical linear mixed models are mostly used. In this study different models were used for varies reasons. Quantile regression was used, because it is more flexible than classical linear mixed models approach. It offers a way to model at any quantile of the response distribution instead of the mean. In addition to quantile regression, Bayesian approach was used since it can allow the use of prior information (informative or non-informative). It can model the quantities which have complex analytical derivation, it is easier to account for other distribution other than the general family of distribution. In the Pharmaceutical industry it is possible to use a small data set in the stability analysis if you have another source of information (prior) whilst in a frequentist this is not possible. Stability quantities can be calculated based on the estimated values from the fitted models.

## 2.1 Linear mixed model

The stability studies are mostly performed to show the degradation pattern of drug product as a function of time. As proposed by Verbeke and Molenberghs, (2000), the linear mixed model has become the main parametric tool for the analysis of such kind of data (continuous longitudinal data). In general, the linear mixed model have to satisfy the following equation.

$$\mathbf{Y_{ijk}} = X_{ij}\beta + Z_{ij}\mathbf{b_i} + \epsilon_{ijk} \tag{1}$$

with: $\mathbf{b_i} \sim N(\mathbf{0}, \mathbf{D})$, $\epsilon_{ijk} \sim N(\mathbf{0}, \Sigma_\mathbf{i})$, $\epsilon_{ijk}$ *and* $\mathbf{b_i}$ *are independent*
where $y_{ijk}$ is the assay result (percent of label claim) from the $j^{th}$ condition(j=1,2) at the $k^{th}$ time point (k =0,1,3,6,12,18 and 24) of the $i^{th}$ batch(i= 1,...,30 for $k > 0$).

$X_{ij}$ is the design matrices for the fixed effects, $Z_{ij}$ is the design matrices for the random effects of known covariates, $\beta$ is the vector containing the parameter estimates for of the fixed effect, $\epsilon_{ijk}$ is the vector containing the residual components, $b_i$ is the vector containing the lots specific deviation from $\beta$. (Verbeke and Molenberghs, 2000). Preliminary, linear mixed model

may contain many fixed effects, and complex covariance structure. In that case, likelihood ratio(LR) test can be applied to select the most parsimonious model.

**Likelihood ratio (LR) test**

Likelihood ratio tests is classical statistical test for comparing nested models with different mean, as well as covariance structures. In case of comparing different mean structures, the LR test is performed using maximum likelihood (ML) estimation method, whereas the tests concerning the covariances structures are performed under restricted maximum likelihood estimation method mixtures of $\chi^2$ with the same weights (0.5) are recommended, otherwise the null hypothesis would be accepted too often, resulting in incorrect simplifying the covariance structure of the model (Verbeke and Molenberghs, 2000).

### 2.1.1 Expiry/Shelf Life

Shelf life being the time at which the drug characteristic remains within an approved specification after manufacture. Mathematically, shelf life ($T_s$) is the abscissa of the intersection of a horizontal line at the specification limit with the confidence or prediction band.
In case of degradation, the lower one-sided 95 percent confidence/prediction limit should be compared to the lower specification limit (LSL). For an attribute known to increase with time, the upper one-sided 95 percent confidence limit should be compared to the upper specification limit (USL). For an attribute whose unknown direction of change with time or non monotone trend, two-sided 95 percent confidence/prediction limits should be calculated and compared to the ULS and LSL (FDA,2004). According to Chow and Shao (1991), the smaller root of the following quadratic equations, if it exists, will be the estimated shelf-life ($T_s$) for condition j:

$$LSL_j = A + B_j * T_s - c * \sqrt{V_s} \quad \Rightarrow \quad (LSL_j - (A + B_j * T_s)^2)^2 = c^2 * V_s \tag{2}$$

$$USL_j = A + B_j * T_s + c * \sqrt{V_s} \quad \Rightarrow \quad (USL_j - (A + B_j * T_s)^2)^2 = c^2 * V_s \tag{3}$$

where A is the fixed intercept and $B_j$ is and the condition-specific slope, respectively, Vs is the variance term for the desired bound and model type, c = $Z_\alpha$. The variance Terms $V_s$ used in

Shelf Life Estimation is defined as follow:

$$V_s = \begin{cases} Var(A + B_j * T_s) + \sigma_b^2 & for \quad confidence \quad interval \\ Var(A + B_j * T_s) + \sigma_b^2 + \sigma_\epsilon^2 & for \quad prediction \quad interval \end{cases}$$

### 2.1.2 Release Limit

Release limits are the intervals bounds formed on the basis of given specifications and real time stability data so that a batch whose true mean at time of manufacture falls within these limits has a high level of assurance ( its mean will remain within specifications throughout shelf life) Based on a 100*(1-$\alpha$)% prediction(confidence) bound on the slope of $j^{th}$ condition($B_j$) at a shelf life (Ts) and a lower(LSL) or upper specification limit(USL). The lower(LRL)/upper(URL) release limit are calculated as follow:

$LRL_j = LSL_j - (B_j * T_s) + c * \sqrt{V_s} \quad and \quad URL_j = USL_j - c * \sqrt{V_\epsilon}$

where

$$V_s = \begin{cases} Var(B_j * T_s) & for \quad confidence \quad interval \\ Var(B_j * T_s) + \sigma_\epsilon^2 & for \quad prediction \quad interval \end{cases}$$

,

## 2.2 Linear quantile mixed model (LQMM)

Linear quantile mixed models (Geraci and Bottai 2014) represent a flexible statistical tool to analyze data from sampling designs such as multilevel, spatial and longitudinal, which induce some form of clustering. Geraci and Bottai (2007) proposed LQMM for longitudinal data using the Asymmetric Laplace distribution(ALD) to model the $\tau^{th}$ conditional quantile of a continuous response variable. They assumed the following regression model:

$$y_{ijk} = \mu_{ij}^{(\tau)} + \epsilon_{ijk}^{(\tau)} \tag{4}$$

where:

$\mu_{ij}^{(\tau)} = X_{ij}\beta^{(\tau)} + Z_{ij}\mathbf{b_i}, \quad \epsilon_{ijk} \sim AL(0, \sigma, \tau), \quad \mathbf{b_i} \sim P(bi|D^\tau)$, and

$\epsilon_{ijk}$ *and* $\mathbf{b_i}$ *are independent.*

Even if the conditional distribution $F_{y_{ijk}|b_i}$ is assumed to be unknown, its $\tau^{th}$ quantile is conveniently estimated as the location parameter $\mu_{ij}^{\tau}$ of an ALD with scale $\sigma^{\tau}$ at a given skewness $\tau$. The skew parameter $\tau$ is set a priori and defines the quantile level to be estimated. Given residual scale parameter($\sigma$) residual variance of y at $\tau^{th}$ quantile is given by

$$\sigma_{\epsilon}^2 = \frac{\sigma^2 * (1 - 2 * \tau + 2 * \tau^2)}{(1 - \tau)^2 * \tau^2} \tag{5}$$

### 2.3 Linear mixed model(Bayesian approach)

Bayesian linear mixed model(BLMM) can be fitted by providing all parameters of the linear mixed model from equation (1), a prior distribution. BLMM with the classical normal distributional assumptions is given by the following:

Level 1: $y_{ij}|\beta, b_i, \sigma^2 \sim N(x_{ij}^T\beta + z_{ij}^T b_i)$
Level 2: $\beta|b_i \sim N(0, D)$
Priors: $\sigma^2 \sim p(\sigma^2), \beta \sim p(\beta)$ and $D \sim p(D)$

The joint posterior distribution for the BLMM is then given by

$$p(\beta, D, \sigma^2, b_1, ...b_n, |y_1, ...y_n) \propto \prod_{i=1}^{n}\prod_{j=1}^{mi} p(y_{ij}|\beta, b_i, \sigma^2) \prod_{i=1}^{n} p(b_i|D)p(\beta)p(D)p(\sigma^2) \tag{6}$$

Now and onward, vague priors will be considered, since there was no historical information provided for the parameters. In a frequentist approach, the random effects are integrated out giving closed expressions for the LMM. The marginal likelihood is then maximized to find the MLEs. In the Bayesian approach, the random effects are sampled together with the fixed-effects parameters. Marginal inference on the fixed effects parameters is obtained by forgetting the sampled random effects parameters(Lesaffre and Andrew, 2012).

#### Producer and consumer risk

When a decision is to be made about a shelf life of a product, the following can be happen, firstly a probability that the product still meeting the quality requirements(good product) will

be falsely determined as violating the specification limit. This probability is called manufacturer/producer risk . Secondly a probability that the product quality violating the specification limit(defective product) at release time will be falsely accepted as conforming at the end of shelf life (consumer risk). Table 1 illustrate these definitions in a good manner(Ilya K. et al, 2011).

Table 1: Consumer and producer risk

| Consumer decision | State of nature | |
|---|---|---|
| | Product is good | Product is defective |
| accept lot | ok | Consumer risk |
| not accept lot | Producer risk | ok |

Consumer and producer risk can be calculated in bayesian framework, by allocating predicted values into categories of (non)acceptable batches at both release and end of shelf life given specifications and release limits.

# 3 Result

## 3.1 Exploratory Data Analysis

As a critical first step in discovering patterns of systematic variation, as well as aspects of random variation that distinguish lots and the implication on model building, graphical techniques were used to explore the individual lots, the mean structure and the covariance structures. Some summary statistics were also used to get more insight in the data.

Table 2 shows the mean and variance of each condition within API at each time point. From this table one can clearly see that the assay amount in each condition decreases with time, except at month 6 and 18 in condition 25C /60%RH, but the increment were so small. They were no observation at time 0 from condition 30C/75%RH on both ingredients because the initial values from that condition were the same as the initials obtained from 25C /60%RH condition. There were small variability between time points. This have to be taken into account during the model building.

Table 2: Summary statisticts by API condition and time

| | API(A) | | API (B) | |
|---|---|---|---|---|
| condition | 25C /60%RH | 30C/75%RH | 25C /60%RH | 30C/75%RH |
| Time | Mean(var) | Mean(var) | Mean(var) | Mean(var) |
| 0 | 98.031(1.0649) | - | 101.441(0.657) | - |
| **1** | 97.976(1.118) | 98.218(0.758) | 101.292(1.176) | 101.624(0.818) |
| **3** | 97.792(1.183) | 97.891(1.256) | 101.31(1.036) | 101.417(0.797) |
| **6** | 97.931(0.499) | 97.876(0.976) | 101.019(0.727) | 100.924(0.910) |
| **12** | 97.529(0.63014) | 97.535(1.005) | 100.3(0.632) | 100.330(1.096) |
| **18** | 97.688(1.4474) | 97.175(0.668) | 99.563(1.045) | 99.781(0.796) |
| **24** | 97.256(0.51064) | 96.813(0.788) | 98.752(0.871) | 99.344(1.262) |

In Figure 1, we explore the individual profiles (black lines) and their average evolution (red line) from the different API and conditions. The first row ( top left and top right panel) represents the API A at storage condition 25C /60%RH and 30C/75%RH respectively.

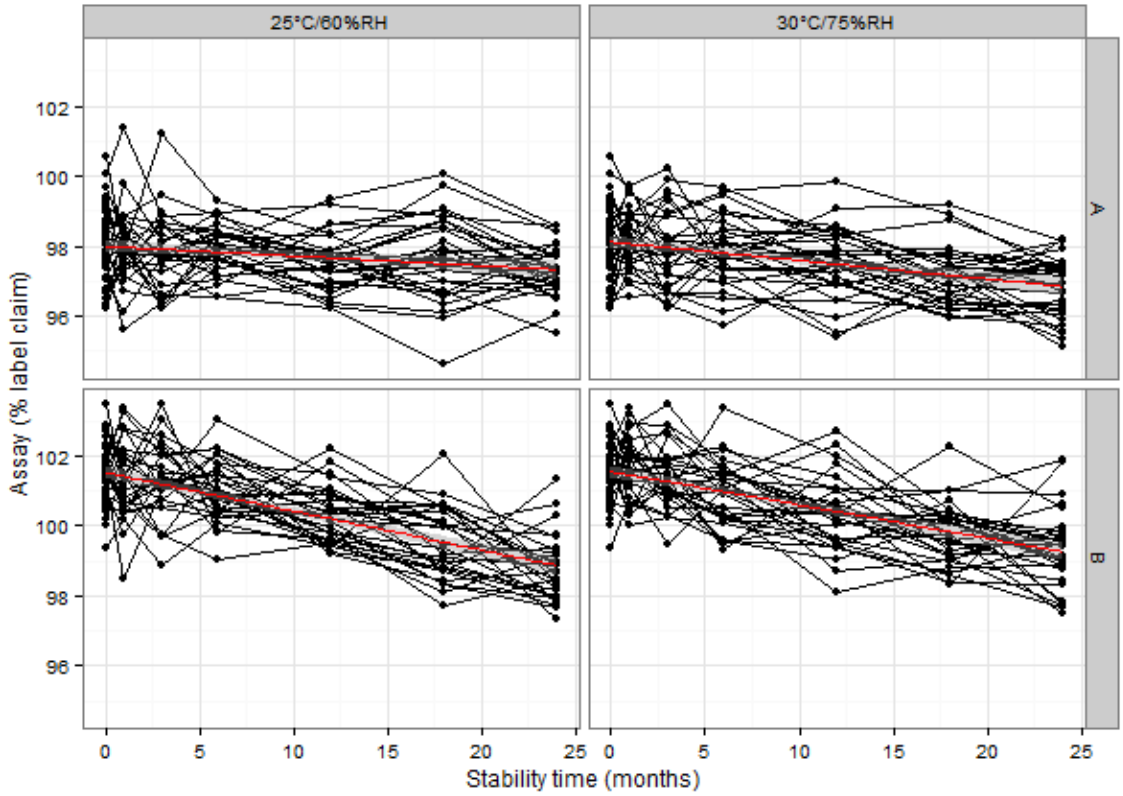
It was observed that the degradation from different lots were more or less similar. The assay from distinct lots show some variability at the beginning. Lots in condition 30C/75%RH have more degradation compared to lots in condition 25C /60%RH on average.

The second row represents the API B (bottom left and bottom right panels) in both storage conditions. Both panels show a clear decreasing as time goes.

In general, the degradation obtained from API B seems to be much faster than that from API A. In all four sub-figures, the average evolution shows a linear trend. As a starting point, model with linear mean structure can be used to study evolution of the assay.

The variance profiles shown in Figure 5 (in appendix), shows a constant variance over time. This suggests that for each API, a model with common constant variances for both conditions seemed to be plausible.

Figure 1: Individual and mean profiles




Red line stands for the mean profile. Black lines stand for individual profiles.

In summary, from the above exploration (summary statistics and plots), for each APIs, model with linear mean structure, with random intercepts and common variance for both conditions seemed to be reasonable. But formal test will be performed to confirm these suggestions.

### 3.2 Model Reduction

In this part the interest was to reduce the mean as well as covariance structures, yielding more parsimonious models. According to Verbeke and Molenberghs (2000), it is sensible strategy to start with a saturated model for the mean structure, as we are dealing with balanced dataset with storage condition labels as the only covariates. A model with both random intercept and slope was chosen as the starting point. Likelihood Ratio(LR) test was conducted to compare the nested models($M_i vs M_j$). For model $M_j$ nested in model $M_i$, the corresponding hypothesis are stated like:

$$hypothesis : \begin{cases} H0 : M_i \\ H1 : M_j \end{cases}$$

The candidate models are:

$M_1$ : Model with saturated mean structure, plus random intercept and slope.
$M_2$ : Model with saturated mean structure, plus a random slope.
$M_3$ : Model with saturated mean structure, plus a random intercept.
$M_4$ : Model with saturated mean structure (fixed effects model).
$M_5$ : Model with a linear mean structure, plus a random intercept.

Table 3 summarizes inferences about the variance components. Different models with the same (saturated) mean structure were fitted based on restricted maximum likelihood estimation method. The p_values were obtained using mixture of chi square($\chi^2$). The test showed that model 3 ($M_3$) was plausible. There was no significant loss by excluding random slope in the model containing random intercept.

Table 3: Model Reduction based on variance structure

| | | API=A | | API=B | |
|---|---|---|---|---|---|
| Model | $\chi^2$ test | LR | Pvalue | LR | Pvalue |
| ($M_1$ vs $M_2$) | $\chi^2_{1:2}$ | 124.60 | $< 0.001$ | 43.2 | $< 0.001$ |
| ($M_1$ vs $M_3$) | $\chi^2_{1:2}$ | **2.40** | **0.789** | **0.00** | **1.00** |
| ($M_3$ vs $M_4$) | $\chi^2_{0:1}$ | 162.60 | $< 0.001$ | 64.80 | $< 0.001$ |

Table 4 shows the results from RL test between random intercept model with a saturated mean structure (under the null hypothesis) and with linear mean structure (under the alternative hypothesis). There was no significant difference between both model(p_value=0.735 and 0.807 for API=A and B respectively). All these tests are in agreement with the exploratory data analysis, a random intercept model with linear mean structure seemed to be plausible. This model was considered from now on.

Table 4: Model Reduction based on mean structure

| | | API (A) | | API (B) | |
|---|---|---|---|---|---|
| Model | $\chi^2$ test | LR | Pvalue | LR | Pvalue |
| ($M_3$ vs $M_5$) | $\chi^2_{10}$ | 6.9 | 0.735 | 6.1 | 0.807 |

The selected model can be written as:

$$Y_{ijk} = \beta_0 + b_i + \beta_1 * cond_{i1} * time_{ijk} + \beta_2 * cond_{i2} * time_{ijk} + \epsilon_{ijk}$$

where $y_{ijk}$ is the assay at the $k^{th}$ time point for the $j^{th}$ storage condition and the$i^{th}$ lot, $\beta_0$ is process mean, $b_i$ is random effect for the $i^{th}$ lot, $\beta_1$ is rate of change for the condition 25C /60%RH, $\beta_2$ is rate of change for the condition 30C /75%RH, $time_{ijk}$ is $k^{th}$ time point for the storage condition and the $i^{th}$ lot, $\epsilon_{ijk}$ are the error terms.

## 3.3 Frequenstist approach

### 3.3.1 Classical linear mixed model

Firstly, an analysis using classical linear mixed effects model was performed. Table 14 (in appendix) summarizes the results from linear mixed model for both APIs, all parameter estimates are highly significant (p_value $< .0001$). For API (A), the monthly degradation was slight higher for storage condition 30C/75%RH compared to condition 25C /60%RH, whereas the opposite was observed for API (B). Table 5 shows some parameter estimates and normality test on residual. The normality assumption was not rejected. The estimate at initial (intercept) or mean value at release (stability time=0) for API A and API B were 98.07 and 101.53, respectively. Meaning that on average the batches being manufactured are released with the API at these values.

Table 5: Estimate at initial, model variability and some test for residual

| API | Intercept (Std. Err.) | Lot Variance | Residual Variance | Minimum Residual | Maximum Residual | Normality Test(Pvalue) |
|---|---|---|---|---|---|---|
| A | 98.07 (0.11) | 0.40 | 0.50 | -2.02 | 1.92 | 0.88 |
| B | 101.53(010) | 0.22 | 0.66 | -2.55 | 2.28 | 0.12 |

Table 6 shows the parameter estimate and the one-sided 95% lower confidence/prediction limit of the change from initial at 12, 18 , 24 and 36 months storage time by API. The annual degradation rate (change from initial at 12 months) in the API A stored at condition 25C/60%RH and 30C/75%RH were 0.38 and 0.60 respectively. The annual drop/degradation in the API B stored at condition 25C/60%RH and 30C/75%RH were 1.37 and 1.16 respectively. Assuming the same degradation rate after follow up time(24 months), the degradation rate after 36 months in API A for drug product stored at condition 25C/60%RH were 1.14 and 1.79 for condition 30C/75%RH. For API B the estimated degradation rate were 4.10 and 3.49 respectively. The rate of degradation after two years (24 months) in both API by condition were also presented in the same table.

Table 6: Estimate with standard error in parenthesis and one sided 95% lower confidence/prediction limit of the change from initial at storage times 12, 18, 24 and 36 months by API

| | | 25C/60%RH | | | 30C/75%RH | | |
|---|---|---|---|---|---|---|---|
| API | Time (months) | Estimate (Std. Err.) | 95% LCI | 95% LPI | Estimate (Std. Err.) | 95% LCI | 95% LPI |
| A | 12 | -0.38 (0.06) | -1.42 | -1.94 | -0.60 (0.06) | -1.64 | -2.15 |
| | 18 | -0.57 (0.09) | -1.61 | -2.13 | -0.89 (0.09) | -1.93 | -2.45 |
| | 24 | -0.76 ( 0.12) | -1.80 | -2.32 | -1.19 (0.12) | -2.23 | -2.75 |
| | 36 | -1.14 (0.18) | -2.18 | -2.70 | -1.79 (0.18) | -2.83 | -3.35 |
| B | 12 | -1.37 (0.07) | -2.14 | -2.91 | -1.16 (0.07) | -1.94 | -2.71 |
| | 18 | -2.05 (0.10) | -2.83 | -3.60 | -1.74 (0.10) | -2.52 | -3.29 |
| | 24 | -2.74 (0.14) | -3.51 | -4.28 | -2.33 (0.14) | -3.10 | -3.87 |
| | 36 | -4.10 (0.21) | -4.88 | -5.65 | -3.49 (0.21 | -4.27 ) | -5.03 |

LPI and LCI are the Lower prediction and confidence limit respectively

Shelf life has been calculated using the lower specification limit(LSL), since the both APIs within drug products under study degrade as time goes. The shelf life can not be more than twice the follow up time but can not exceed the follow up time plus 12 months (Shein-chung et al 2007). In this study the shelf life can not exceed 36 months. Based on FDA requirements, the shelf life of drug product is the minimum shelf life between it's APIs.

Table 7 shows the calculated shelf lives based on the lower 95% confidence/prediction limit. At storage condition 25C/60%RH, the minimum shelf life based on both confidence and prediction interval between both APIs was 43 months which is greater than the desired shelf life(36 months). For this condition, it is possible to extend it to 36 months. The same conclusion can be drawn for the shelf life based on confidence interval for condition 30C/75%RH. But for 30C/75%RH, the shelf life based on prediction interval can not be extended to 36 months, since its minimum shelf life was 30 months.

Table 7: Shelf life based on the lower 95% confidence/prediction limit

| API | Shelf life based on the 95% LCI | | Shelf life based on the 95% LPI | |
|---|---|---|---|---|
| | 25C/60%RH | 30C/75%RH | 25C/60%RH | 30C/75%RH |
| A | 60 | 40 | 46 | 30 |
| B | 49 | 57 | 43 | 51 |

LPI and LCI are the Lower prediction and confidence limit respectively

Table 8: Estimate of the release limits based on the one sided 95% lower confidence/prediction considering a shelf life of 24 and 36 months by API

| | | 25C/60%RH | | 30C/75%RH | |
|---|---|---|---|---|---|
| API | Time (months) | 95% LCI | 95% LPI | 95% LCI | 95% LPI |
| A | 24 | 95.96 | 96.39 | 96.94 | 97.37 |
| | 36 | 96.44 | 97.09 | 97.34 | 97.99 |
| B | 24 | 97.55 | 97.96 | 98.68 | 99.09 |
| | 36 | 98.83 | 99.44 | 99.87 | 100.48 |

LPI and LCI are the Lower prediction and confidence limit respectively

Release limits based on the CI will relate to the batch mean, but for better assurance it is recommended to consider the release limits based on the prediction interval as this relates to the individual tablets (accounting for residual variability) and this can pe used to predict future batches under the same circumstance. Table 8 shows the release limits by APIs and condition at the end of current shelf life(24 months) and 36 months. Given the shelf life of 24 months, the release limit for API A ( for API B) on the drug product stored at condition 25C/60%RH was 96.39 (resp. 97.96). For condition 30C/75%RH, the estimated release limit was 97.37(resp. 99.09) for API A ( for API B). For a shelf life of 36 months, the release limit was 97.09(resp. 99.44) for API A ( for API B) for the drug product stored at condition 25C/60%RH. For condition 30C/75%RH, the estimated release limit was 97.99(resp. 100.48) for API A ( for API B). Thus the amount of ingredient to put in the drug product in order to ensure the quality at

the end of shelf life, depends on API as well as storage condition.

### 3.4 Linear quantile mixed model(LQMM)

In this part, the interest was to fit linear quantile mixed model at $0.05^{th}$. This quantile can correspond to the lower confidence interval of the classical linear mixed model. The basic idea was to offer more protection and reduce the risk to the consumer. Table 15 (in appendix) shows the result from the fitted linear quantile mixed model $0.05^{th}$ quantile by APIs and storage condition. Most of the parameter estimates of both models were significant.

The estimate at initial, estimated between lot, and within lot (residual) variability are shown in Table 9. The $0.05^{th}$ quantile at release is 97.63 and 101.21 for API A and API B respectively. At release at least 95% of the distribution of assay values for API A and API B exceeded these values. At $0.05^{th}$ quantile, the estimated variability between lots was greater than the within lots variability. In other words, at $0.05^{th}$ quantile, most of the variability in the assay distribution not explained by the mean structure was explained by unobserved heterogeneity between lots.

Table 9: Estimate at initial, estimated lot variance, estimated residual variance

| API | Estimate at initial | Lot Variance | Residual Variance |
|---|---|---|---|
| A | 97.63 | 1.04 | 0.07 |
| B | 101.21 | 0.99 | 0.07 |

Table 10 shows the estimates at $0.05^{th}$ quantile and their one-sided 95%lower prediction limit of the change from initial at 12 (annual degradation rate), 18 , 24 and 36 months storage time by API. At $0.05^{th}$ quantile, the degradation rate in API A and B, at the end of current shelf life(2 years) was 0.58 and 3.21 respectively for drug product stored at 25C/60%RH, and 1.19 and 2.71 for storage condition 30C/75%RH.

Table 10: Estimate with standard error and one sided 95% lower prediction limit of the change from initial at storage times 12, 18, 24 and 36 months by API

| API | Time (months) | 25C/60%RH Estimate (Std. Err.) | 25C/60%RH 95% LPI | 30C/75%RH Estimate (Std. Err.) | 30C/75%RH 95 % LPI |
|---|---|---|---|---|---|
| A | 12 | -029(0.04) | -0.29 | -0.59(0.05) | -0.6 |
| | 18 | -0.43(0.9) | -0.44 | -0.89(0.12) | -0.9 |
| | 24 | -0.58(0.15) | -0.59 | -1.19(0.21) | -1.2 |
| | 36 | -0.87(0.34) | -0.88 | -1.78(0.48) | -1.8 |
| B | 12 | -1.61(0.02) | -1.61 | -1.36(0.04) | -1.37 |
| | 18 | -2.41(0.05) | -2.41 | -2.04(0.09) | -2.05 |
| | 24 | -3.21(0.09) | -3.22 | -2.71(0.17) | -2.73 |
| | 36 | -4.82(0.20) | -4.83 | -4.08(0.38) | -4.1 |

LPI and LCI are the Lower prediction and confidence limit respectively

Table 11 shows the estimated shelf lives results at the $0.05^{th}$ quantile. The smallest shelf life for the drug product stored at condition 30C/75%RH was 30 months. Thus we can not extend shelf life to three years. At storage condition 25C/60%RH the smallest shelf life was greater than 36 months, then it was possible to extend the shelf life.

Table 11: Shelf life for both APIs

| API | 25C/60%RH | 30C/75%RH |
|---|---|---|
| A | 47 | 30 |
| B | 39 | 50 |

Table 12 shows the estimated release limit at the end of 24 and 36 months by API and storage condition. It shows how much ingredients (API A and B) required at release time to maintained the quality of the drug product until the end of 24 months as well as 36 months.

Table 12: Release limit for both APIs

| API | Time (months) | 25C/60%RH | 30C/75%RH |
|---|---|---|---|
| A | 24 | 96.1 | 96.3 |
| | 36 | 96.65 | 96.95 |
| B | 24 | 98.29 | 97.72 |
| | 36 | 99.82 | 99.08 |

### 3.4.1 Bayesian approach

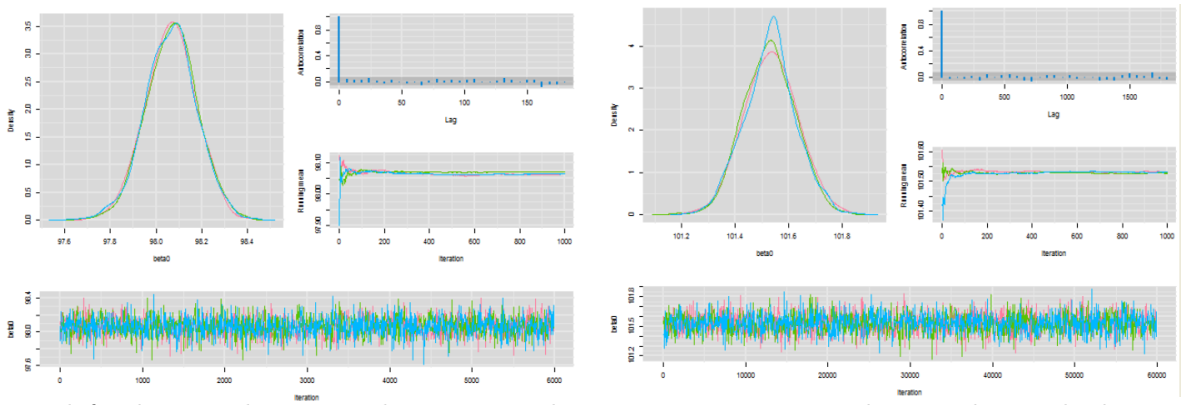
In this part the linear mixed model with only random intercept was considered. Before interpretation, convergence have been assessed to ensure how closer we are to the true posterior distribution. Diagnostics tools such as trace plot, Running mean and autocorrelation plots and formal test like Raftery-Lewis, The Gelman-Rubin and crude measure of effective sample size for each parameter were used.

Three chains , each with 120,000 iteration were run, half of them were discarded(burn-in), to maintain independence between iterations. 60 thinning was used. Trace , autocorrelation and density plots were used to assess convergence. All Diagnostic plots for the regression parameter estimates on both API are displayed in Figure 2 to Figure 4. They indicated that the sampling was done in almost independent manner. The chains did not depend on their initial values and stationarity was achieved. in general all plots showed excellent convergence and good mixing rate. The same conclusions can be drawn for other parameters presented in Figure 6 to Figure 13 (in appendix).

Figure 2: Diagnostic plots for intercept of both API

(a) intercept for API A

(b) intercept for API B

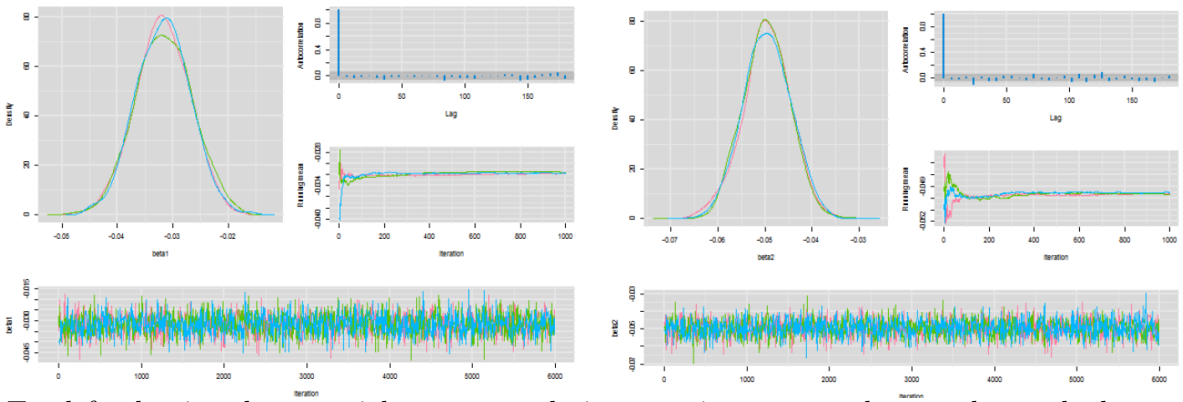


Top left: density plot, top right: autocorrelation, running mean and trace plot at the bottom

Figure 3: Diagnostic plots for both slopes of API A

(a) Condion 25C /60%RH ($\beta_1$)

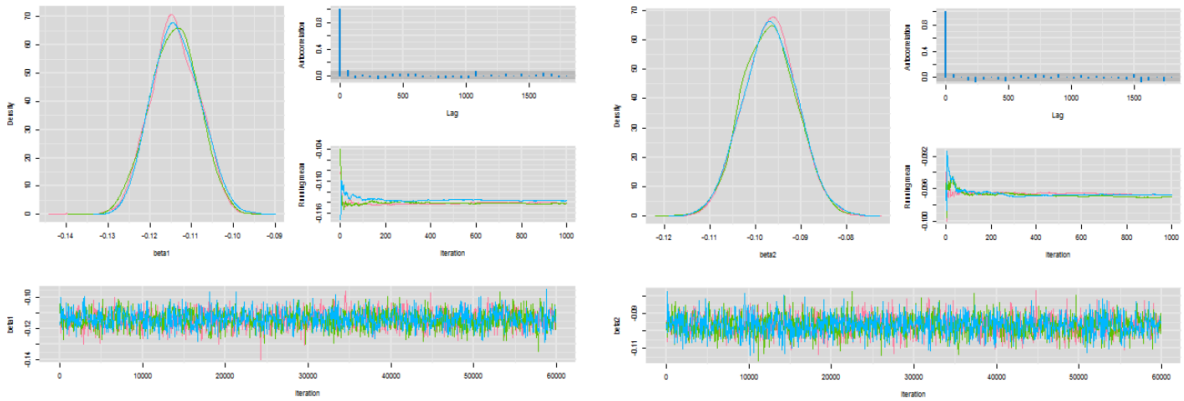
(b) Condion 30C /75%RH ($\beta_2$)



Top left: density plot, top right: autocorrelation, running mean and trace plot at the bottom

Figure 4: Diagnostic plot for both slopes of API B

(a) Condion 25C /60%RH ($\beta_1$) (b) Condion 30C /75%RH ($\beta_2$)



Top left: density plot, top right: autocorrelation, running mean and trace plot at the bottom

The convergence results are shown in Table 16 (in appendix). Potential scale reduction factor (Psrf) and its upper 97.5% CI for all parameters were more or less one, according to Brooks-Gelman-Rubin (BGR) convergence to posterior distribution was achieved. This conclusion was valid to both API. Independence of iterations was also confirmed by the obtained effective sample size which was almost the same as the length of the chains (3000). The smallest effect samples size was half of the total length of the chains, which was not bad for independence of the iterations. According to Raftey-lewis (RL) diagnostic, the dependence factor greater than five is an indication of problems such bad starting values and high autocorrelation and so on. For all the parameters, the dependence factors were less than five. The last diagnostic test used in this study was the Geweke diagnostic test reported in Table 17. It is univariate test which checks the convergence for each chain independently. From Table 16 we can see that almost all parameter estimates are less than 1.96 in absolute value.

Based on informal and formal tests, we can conclude that, the burn-in of 60000 was enough to forget the initial values, there were no dependence of iterations, stationarity and higher mixing rate were attained. In other words, we can say that most probably, the estimates were derived from the true posterior distribution.

Table 18 summarizes the result from bayesian linear mixed model ( all parameter estimates with their standard deviation in bracket, 95% lower credible interval). Vague priors were used. This approach yields parameters estimates close to those from classical linear mixed model. Hence change from initial time up to 12, 18, 24, 36 months are more or less similar in magnitude as the corresponding change from frequentist approach. It was observed that the shelf life for drug product can be extend to 36 months if it was stored at condition 25C/60%RH but not at condition 30C/75%RH. Release limit are almost similar to the those computed based on parameter estimates from classical linear mixed model. Bayesian approach offer a way to obtain the consumer and producer risk.

The estimated producer and consumer risk estimates are presented in Table 13. For both shelf lives, the risks of rejecting a good drug product at release time were almost zero. In most of the case, the consumer risk were more or less zero, excerpt for drug product stored at condition 30C/75%RH with 36 months of shelf life. This was not surprising because at this condition, it was not possible to extend the shelf life to 36 months.

Table 13: Producer and consumer risk based on 24 and 36 months

| | | | 25C/60%RH | | 30C/75%RH | |
|---|---|---|---|---|---|---|
| API | Time (months) | Risk | Estimate | 95% LCI | Estimate | 95% LCI |
| A | 24 | Producer | 0.00 | 0.00 | 0.00 | 0.00 |
| | | Consumer | 0.01 | 0.00 | 0.02 | 0.01 |
| | 36 | Producer | 0.00 | 0.00 | 0.00 | 0.00 |
| | | Consumer | 0.02 | 0.01 | 0.09 | 0.05 |
| B | 24 | Producer | 0.00 | 0.00 | 0.00 | 0.00 |
| | | Consumer | 0.00 | 0.00 | 0.00 | 0.00 |
| | 36 | Producer | 0.00 | 0.00 | 0.00 | 0.00 |
| | | Consumer | 0.01 | 0.00 | 0.01 | 0.00 |

LCI is the lower credible interval.

# 4 Discussion and Conclusion

In this project, we aimed at assessing different statistical models in quantifying pharmaceutical stability, to calculate shelf life , release limit given shelf life, the change in annual rate, change from initial and quantifying the producer and/or consumer risk. To check if it is possible to extend the shelf life from 24 to 36 months. Different statistical models, among which linear mixed models, quantile regression (using both frequentist and bayesian approach), were applied. The exploratory data analysis suggested a linear mean structure model with random intercepts. The model selection was done based on likelihood ratio test under (restricted) maximum likelihood estimation method, and the model from the exploratory analysis was confirmed.

For all regression parameters, frequentist and bayesian approaches had almost the same estimated values, since vague priors were used for the latter approach . Although the slopes were significantly different from zero, their magnitudes were very small. The highest degradation rate(-3.21) after two years was from API B at storage condition 25C /60%RH. Lower prediction bound was used in calculation of shelf life, release limit and producer/ consumer risk, since the assay of a drug under study decreased with time.

If shelf life of API A and API B are XX, and XY respectively with FT as follow up time(FT=24 month is this study). So the shelf life for drug X was set as min(XX, XY, 2FT , FT+12). This mean that in this study, it can never exceed 36 months as per regulations. For the drug product stored at 25C /60%RH and 30C/75%RH, the shelf life was 36 and 30 months respectively. Thus it was not possible to extend the shelf life for drug product stored at condition 30C/75%RH.

For drug product stored at condition 25C/60%RH , if you want to have a shelf life of 24 months you have to release your batches at mean value of 95.96 and 97.55, and your tablets at values 96.39 and 97.96 for API A and B respectively. For condition 30C/75%RH you have to release your batches mean at 96.94 for API A and 98.68 for API B, and your tablets at values 97.37 for API A and 99.09 for API B. For a shelf life of 36 months the release limit for batches mean was 96.44 and 98.83 for API A and B, but 97.09 and 99.44 for API A and B, on individual tablets for drug product stored at condition 25C/60%RH. The same analysi for condition 30C/75%RH. The producer and consumer risk were very small, except the consumer risk for a drug product stored at condition 30C/75%RH. The quantile regression was used in

order to offer more protection to the consumer, but based on the results obtained from quantile mixed model the same conclusion as classical linear mixed models was drawn.

For future research, it will be better to account for the correlation between APIs by using joint models, to quantify the pharmaceutical quantity. using tolerance intervals and to fit bayesian quantile mixed model.

## References


[1] Allen, P., Dukes, G., and Gerger, M. (1991). Determination of Release Limits: A General Methodology. *Pharmaceutical Research* .27 (3). p. 1210-1213

[2] FDA. (1997). Guidance for Industry: Dissolution Testing of Immediate Release Solid Oral Dosage Forms. *The United States Food and Drug Administration*, Rockville, MD.

[3] FDA. (1998). Guidance for Industry: Stability Testing of Drug Substances and Drug Products (draft guidance). *The United States Food and Drug Administration*, Rockville, MD.

[4] FDA. (2004). Guidance for Industry: Q1E Evaluation of Stability Data. *The United States Food and Drug Administration*, Rockville, MD.

[5] Geraci, M. and Bottai, M. (2007). Quantile regression for longitudinal data using the asymmetric Laplace distribution. *Biostatistics*. 8. p. 140-154

[6] Geraci, M. and Bottai, M. (2014). Linear quantile mixed models. *Biostatistics*. 20(2). p. 461-479

[7] Hsin-ya, L., Pao-chu W., and Yung-jin L. (2010). An R package for drug stability data analysis. *Computer methods and program in biomedicine*. 24(2). p. 140-148.

[8] Ilya, K., Ilana, S., Francesca, P. Cathy, B. Ales, F., and Paolo, Z. (2011). Long-term stability study of drug products and out-of-specification test results. *General paper*. 29(7) p. 615-622

[9] Lesaffre, E. and Andrew B. L. (2012). *Bayesian Biostatistics*. New-York: John Wiley and Sons

[10] Shein-chung, C., Byron, J., Jen-peu, L., and karl, E. P. (2007). *Statistical design and analysis of stability studies*. Biostatistic series. Broken Sound Parkway:Taylor and Francis group.

[11] Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer Series in Statistics. New-York: Springer

[12] Chow, S.C. and Shao, J. (1991). Estimating drug shelf-life with random batches. *Biometrics*. 47 (3). p. 1071-1079.

# Apendix

Figure 5: Variance profiles

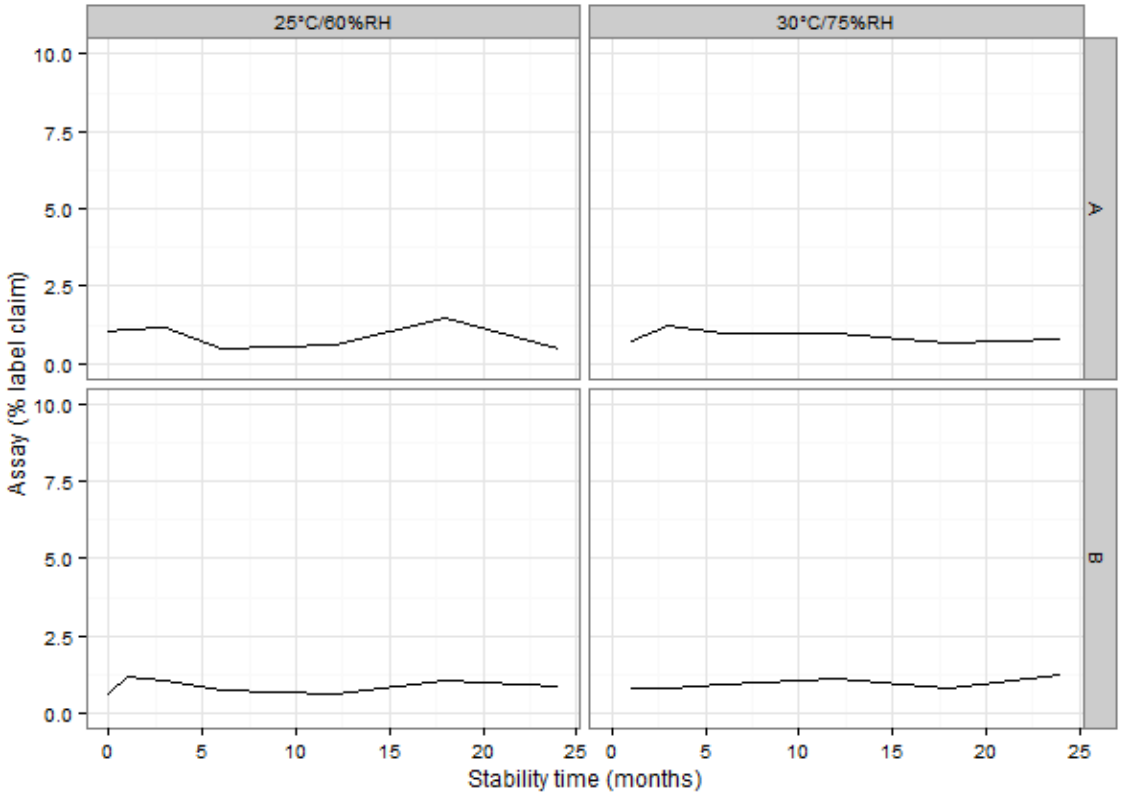


Table 14: LMM frequentist approach

| effects | Parameter | API (A) | | API(B) | |
|---|---|---|---|---|---|
| | | Estimate(STD) | P_value | Estimate(STD) | P_value |
| intercept | $\beta_0$ | 98.071(0.109) | $< .0001$ | 101.532(0.095) | $< .0001$ |
| Cond1*time | $\beta_1$ | -0.032(0.005) | $< .0001$ | -0.114(0.006) | $< .0001$ |
| Cond2*time | $\beta_2$ | -0.050(0.005) | $< .0001$ | -0.097(0.006) | $< .0001$ |
| | | Model variability | | | |
| Between lot | $\sigma_b^2$ | 0.3999 | - | 0.223 | - |
| Within lot | $\sigma^2$ | 0.498 | - | 0.658 | - |

Cond1 and cond2 are condition 25C /60%RH and 30C/75%RH respectively

Table 15: Linear quantile mixed model

| Effects | Parameters | API(A) | | API(A) | |
|---|---|---|---|---|---|
| | | Estimate | P_value | Estimate | P_value |
| Intercept | $\beta_0$ | 97.63 | $< 0.001$ | 101.21 | $< 0.001$ |
| time*cond1 | $\beta_1$ | 0.0 3 | 0.4 | -0.14 | $< 0.001$ |
| time*cond2 | $\beta_2$ | 0.005 | $< 0.001$ | -o.11 | $< 0.001$ |
| | | Model variability | | | |
| Between lot | $\sigma_b$ | 1.04 | | 1.07 | |
| Within lot | $\sigma_\epsilon$ | 0.7 | | 0.7 | |

Table 16: Convergence assessment

| **Effects** | API (A) | | | API (B) | | |
|---|---|---|---|---|---|---|
| | Df (I)) | Psrf(97.5% quantile | N.Eff | Df (I)) | Psrf(97.5% quantile) | N.Eff |
| intecept | 1.01 | 0.999(1.00) | 3000 | 2.49 | 1.000(1.00) | 3000 |
| Cond1*time | 1.01 | 1.001(1.01) | 3000 | 1.01 | 1.000(1.00) | 3000 |
| Cond2*time | 0.988 | 1.001(1.00) | 3000 | 0.998 | 1.001(1.00) | 3000 |
| Shelf lifeC_Cond1 | 1.1 | 1.008(1.02) | 3000 | 1.06 | 1.002(1.01) | 1800 |
| Shelf lifeC_Cond2 | 2.27 | 1.002(1.01) | 1500 | 1.05 | 1.004(1.01) | 3000 |
| Shelf lifeP_Cond1 | 1.15 | 1.008(1.02) | 3000 | 1.03 | 1.002(1.01) | 1800 |
| Shelf lifeP_Cond2 | 2.45 | 1.002(1.01) | 1500 | 1.05 | 1.004(1.01) | 3000 |
| Based on shelf life of 24 months | | | | | | |
| RLC_Cond1 | 2.27 | 1.000(1.00) | 3000 | 1.00 | 1.009(1.03) | 3000 |
| RlC_Cond2 | 3.21 | 1.000(1.00) | 3000 | 1.02 | 1.009(1.03) | 3000 |
| RLP_Cond1 | 1.18 | 1.000(1.00) | 3000 | 1.15 | 1.009(1.03) | 3000 |
| RLP_Cond2 | 1.25 | 1.000(1.00) | 3000 | 1 08 | 1.009(1.03) | 3000 |
| Producer24 (PR0) | 4.2 | 1.02(1.03) | 3000 | 1.96 | 1.005(1.01) | 3000 |
| Consumer24 (CR1) | 1.28 | 1.001(1.01) | 3000 | 1.04 | 1.000(1.00) | 3000 |
| Consumer24 (CR2) | 1.09 | 1.002(1.01) | 1700 | 4.52 | 1.004(1.01) | 1500 |
| Based on shelf life of 36 months | | | | | | |
| RLC_Cond1 | 3.01 | 1.00(1.00) | 3000 | 1 | 1.00(1.03) | 1700 |
| RlC_Cond2 | 3.21 | 1.00(1.00) | 2800 | 1.02 | 1.00(1.03) | 3000 |
| RLP_Cond1 | 1.18 | 1.00(1.00) | 3000 | 1.15 | 1.00(1.03) | 2000 |
| RLP_Cond2 | 1.25 | 1.00(1.00) | 1600 | 1 08 | 1.00(1.03) | 3000 |
| Producer36 (PR0) | 4.2 | 1.00(1.03) | 3000 | 1.96 | 1.005(1.01) | 3000 |
| Consumer36 (CR1) | 1.28 | 1.00(1.01) | 3000 | 1.04 | 1.000(1.00) | 3000 |
| Consumer36 (CR2) | 1.09 | 1.00(1.01) | 3000 | 4.52 | 1.004(1.01) | 2500 |
| Model variability | | | | | | |
| between lot | 2.55 | 1.001(1.00) | 3000 | 2.67 | 1.014(1.04) | 2300 |
| Residual variance | 1.03 | 1.004(1.01) | 3000 | 1.04 | 0.999(1.00) | 3000 |

RLC and RLP are the release limit based on confidence and prediction interval respectively, cond1 and cond2 are the storage condition 25C/60%RH and 30C/75%RH receptively.

Table 17: Geweke diagnostic

| **Effects** | API A | | | API B | | |
|---|---|---|---|---|---|---|
| | chain1 | chain 2 | chain 3 | chain1 | chain 2 | chai3 |
| intecept | -0.362 | 0.057 | -0.880 | 1.837 | 0.009 | 0.038 |
| Cond1*time | 1.366 | 0.331 | -0.648 | -1.123 | -0.850 | 0.460 |
| Cond2*time | 1.548 | 0.886 | -0.401 | 0.614 | -1.560 | 0.926 |
| Shelf life1 | 1.290 | 0.408 | -1.267 | 0.142 | -0.689 | -0.196 |
| Shelf life2 | 1.585 | 0.340 | -1.217 | 1.088 | -1.193 | 0.391 |
| Based on shelf life of 24 months | | | | | | |
| RLC_Cond1 | 1.04 | -0.41 | -1.29 | 1.12 | 0.19 | -0.25 |
| RLC_Cond2 | 0.19 | -1.55 | -1.02 | -1.58 | -1.20 | -0.21 |
| RLP_Cond1 | 1.67 | 0.40 | 1.95 | 1.74 | 0.79 | -1.32 |
| RLP_Cond2 | -1.24 | -0.09 | -1.73 | 1.47 | 0.64 | 0.14 |
| Consumer R0 | 0.609 | -0.113 | -0.519 | 2.152 | 0.005 | -0.956 |
| Producer R1 | -0.148 | 0.662 | -0.468 | 1.649 | 0.4690 | -0.934 |
| Producer R2 | -0.731 | -0.968 | 1.142 | -1.269 | 0.931 | 0.289 |
| Based on shelf life of 36 months | | | | | | |
| RLC_Cond1 | 1.68 | -0.40 | -1.21 | 1.54 | 0.67 | -0.51 |
| RLC_Cond2 | 2.14 | -0.51 | -1.22 | 1.57 | 0.09 | 1.25 |
| RLP_Cond1 | 1.94 | -0.81 | -1.92 | 1.66 | 0.679 | -0.21 |
| RLP_Cond2 | -0.94 | -1.28 | -1.29 | 1.47 | 1.52 | -1.251 |
| Consumer R0 | 0.609 | -0.113 | -0.519 | 2.152 | 0.005 | -0.956 |
| Producer R1 | -0.148 | 0.662 | -0.468 | 1.649 | 0.4690 | -0.934 |
| Producer R2 | -0.731 | -0.968 | 1.142 | -1.269 | 0.931 | 0.289 |
| Model variability | | | | | | |
| between lot | -1.845 | 0.639 | 0.868 | -0.378 | -0.622 | 0.650 |
| Residual variance | 0.516 | -0.555 | -0.542 | 0.876 | 0.0447 | 1.736 |

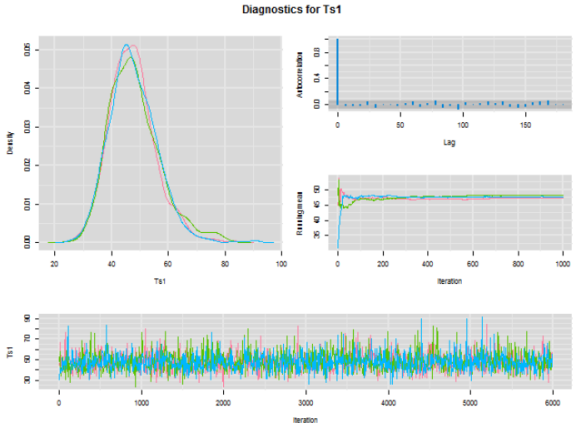
Table 18: LMM bayesian approach

| effects | Parameter | **API (A)** | | **API (B)** | |
|---|---|---|---|---|---|
| | | Estimate(STD) | 95% credible interval | Estimate(STD) | 95% credible interval |
| intecept | $\beta_0$ | 98.071(0.110) | (97.854,98.279) | 101.529(0.100) | (101.333,101.733) |
| Cond1*time | $\beta_1$ | -0.032(0.005) | (-0.042,-0.022) | -0.114(0.006) | (-0.125,-0.103) |
| Cond2*time | $\beta_2$ | -0.050(0.005) | (-0.059,-0.039) | -0.097(0.006) | (-0.108,-0.086) |
| Shelf life_P_Cond1 | TsP1 | 47.578(8.559) | (33.637,66.813) | 43.566(2.247) | (39.418,48.188) |
| Shelf life_P_Cond2 | TsP2 | 29.97(4.024) | (22.574,38.634) | 51.387(2.983) | (45.709,57.405) |
| Shelf life_C_Cond1 | TsC1 | 64.25(11.68) | (44.71,89.77) | 50.23(2.72) | (45.14,55.88) |
| Shelf life_C_Cond2 | TsC2 | 40.41(5.14) | (30.96,51.12) | 59.17(3.61) | (52.56,66.51) |
| | | Release limit given shelf life=24 months | | | |
| RLC_Pond1 | RLP1 | 96.33(0.13) | (96.68, 97.18) | 99.07(0.15) | (98.97,99.37) |
| RLC_Pond2 | RLP2 | 97.36(0.13) | (97.11,97.61) | 98.66(0,14) | (98.39,98.95) |
| RLC_Cond1 | RLC1 | 95.76(0.12) | (95.53 96.00) | 97.73(0.14) | (97.46,98.01) |
| RLC_Cond2 | RLC2 | 96.19(0.12) | (95.95,96.43) | 97.32(0.14) | (97.23,97.59) |
| | | Release limit given shelf life=36 months | | | |
| RL_P_Cond1 | RLP1 | 97.31(0.18) | (96.97, 97.67) | 100.45(0.21) | (100.04,100.87) |
| RL_P_Cond2 | RLP2 | 97.96(0.19) | (97.59, 98.32 ) | 99.83(0.21) | (99.68,100.24) |
| RL_C_Cond1 | RLC1 | 96.15(0.18) | (95.81, 96.49) | 99.11(0.21) | (98.70,99.51) |
| RL_C_Cond2 | RLC2 | 96.79(0.18) | (96.43, 97.15) | 98.(0.21) | (98.08,98.89) |
| | | Model variability | | | |
| Between lot | $\sigma_b^2$ | 0.435 | - | 0.246 | - |
| Within lot | $\sigma^2$ | 0.500 | - | 0.663 | - |

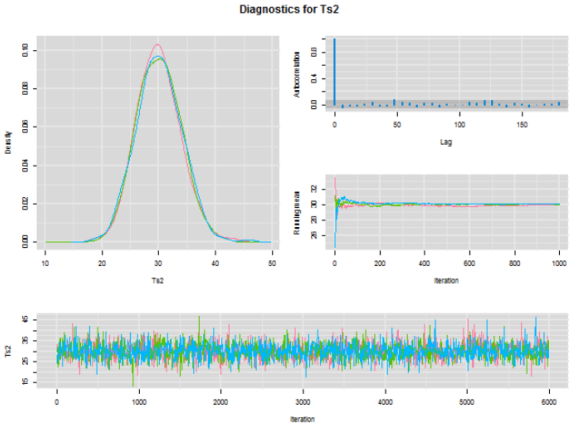
## Diagnostic plot for API A

Figure 6: Diagnostic plots for Shel lives

(a) Condion 25C /60%RH



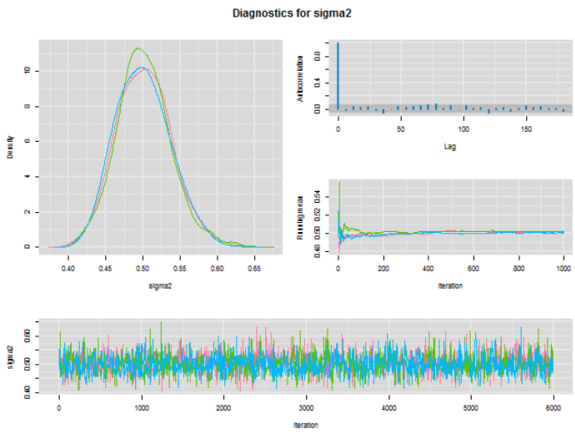

(b) Condion 30C /70%RH



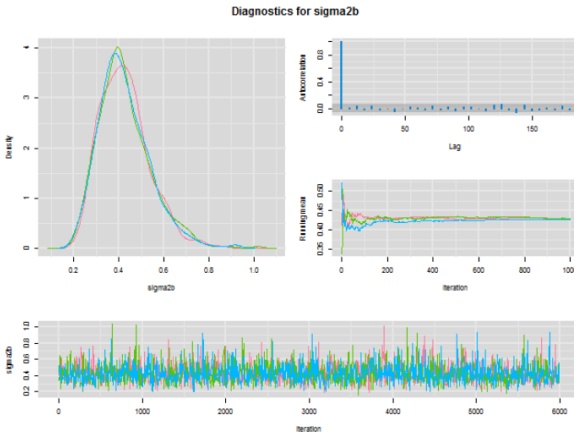

Top left: density plot, top right: autocorrelation, running mean and trace plot at the bottom

Figure 7: Diagnostic plots for variance parameter

(a) Measurement error



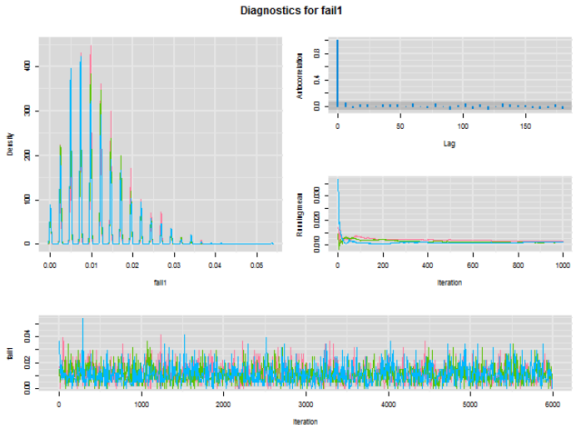

(b) Between lots variability



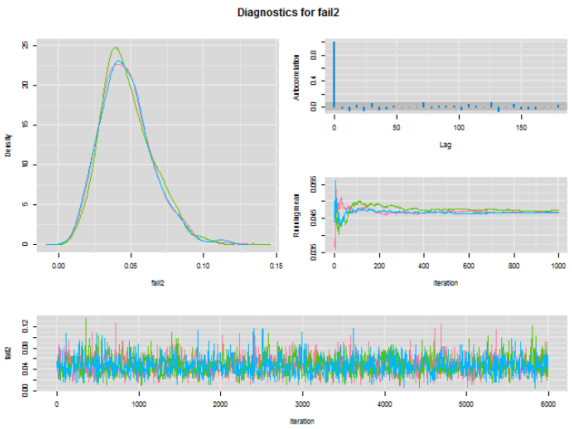

Top left: density plot, top right: autocorrelation, running mean and trace plot at the bottom

Figure 8: Diagnostic plots for producer risk

(a) Condion 25C /60%RH

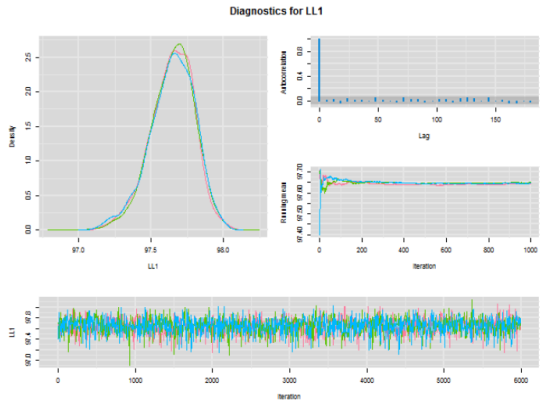



(b) Condion 30C /75%RH




Top left: density plot, top right: autocorrelation, running mean and trace plot at the bottom

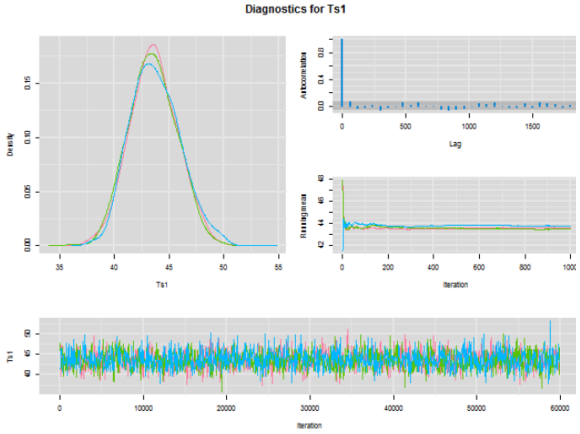
Figure 9: Diagnostic plot for release limit



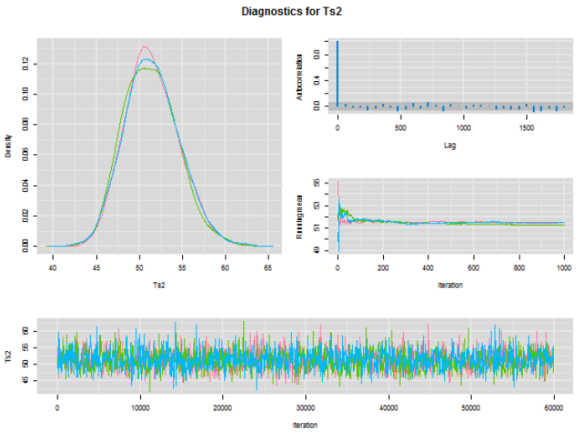
## Diagnostic plot for API B

Figure 10: Diagnostic plots for Shel lives

(a) Condion 25C /60%RH



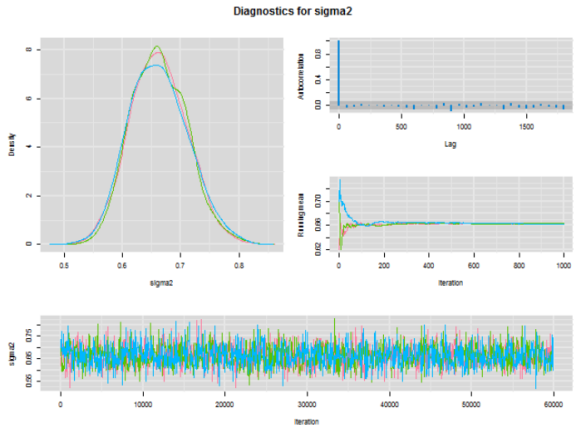

(b) Condion 30C /70%RH



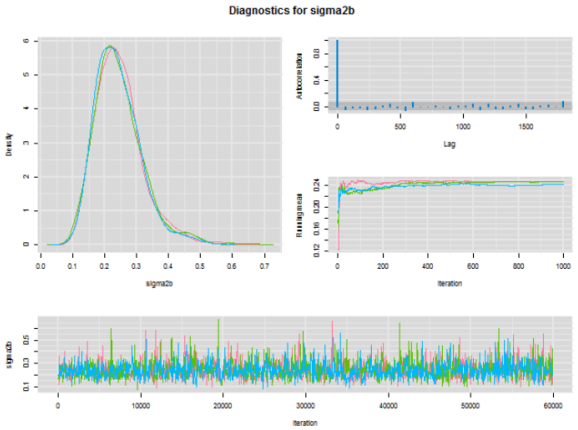

Top left: density plot, top right: autocorrelation, running mean and trace plot at the bottom

Figure 11: Diagnostic plots for variance parameter

(a) Measurement error



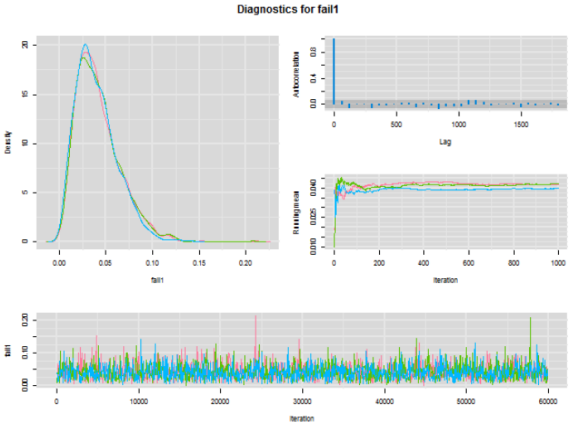

(b) Between lots variability



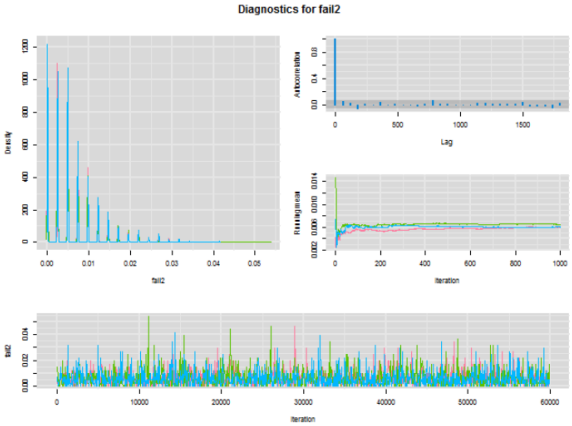

Top left: density plot, top right: autocorrelation, running mean and trace plot at the bottom

Figure 12: Diagnostic plots for producer risk

(a) Condion 25C /60%RH

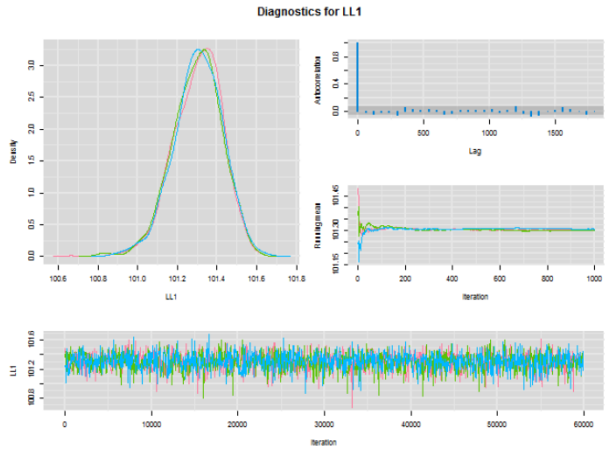



(b) Condion 30C /75%RH




Top left: density plot, top right: autocorrelation, running mean and trace plot at the bottom

Figure 13: Diagnostic plot for release limit



# Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:
**Statistical models for stability studies**

Richting: **Master of Statistics-Biostatistics**
Jaar: **2015**

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,

**Uwimpuhwe, Germaine**

Datum: **1/09/2015**