Estimating the prediction accuracy of production values

Severi Saastamoinen

School of Science

Bachelor's thesis
Espoo 12.4.2021

Supervisor and advisor

Prof. Ahti Salo

Aalto University
School of Science

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of the bachelor's thesis

Author Severi Saastamoinen

Title Estimating the prediction accuracy of production values

Degree programme Engineering Physics and Mathematics

Major Mathematics and Systems Sciences **Code of major** SCI3029

Teacher in charge and advisor Prof. Ahti Salo

Date 12.4.2021 **Number of pages** 19+1 **Language** English

Abstract

In this Bachelor's thesis, we study the effects and behavior of random unexpected delays and malfunctions of machinery in four production facilities in Finland. The purpose of the Thesis was to support the better prediction of the production and help research the behaviour of the production in various circumstances.

The prediction was made with a simple model, which assumes the daily production to be equal throughout the month, excluding days, which were predetermined to be maintenance breaks. The research was done by dividing the monthly prediction for each day, which was supposed to be an active workday and by subtracting the prediction from the actual production on each day. Thus it was possible to compare the accuracy of the prediction daily and calculate the mean, median, moving average, separate categories of days with exceptionally low output and illustrations with confidence intervals for the data.

During most days the production was more, than was predicted, because the prediction was designed to take into account the typical fluctuation of the output. The output was greater than the average in most days, whereas during the days, when the output was lower than the average it was often clearly lower. There were less days, when the output was lower than higher compared to the average, but the large drop in output during the days of lower production explains their great influence to the average production. Due to breaks in production the average output was nearly the same as the prediction and the median of the output was considerably higher. It was worth noting, how unplanned breaks and longer than planned maintenance periods caused significant losses and caused the mean of the output to be well below it's median.

All this means, that during long periods of time the the prediction is very reliable, because the mean of the output is almost the same as the values given by the prediction. During shorter periods of time, however, there can be notable discrepancies between the prediction and the output. The prediction could be adjusted by calculating the average amount of days with which maintenance breaks get extended and adding that amount of days to the length of the breaks in the model. The prediction for that month could be divided to the remaining days, making the prediction of each day closer to the output.

Keywords Production, Prediction, Accuracy, Statistics, Time series analysis

Aalto University
School of Science

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi
Tekniikan kandidaatintyön tiivistelmä

Tekijä Severi Saastamoinen

Työn nimi Estimating the prediction accuracy of production values

Koulutusohjelma Matematiikka ja teknillinen fysiikka

Pääaine Matematiikka ja systeemianalyysi **Pääaineen koodi** SCI3029

Vastuuopettaja ja ohjaaja Prof. Ahti Salo

Päivämäärä 12.4.2021 **Sivumäärä** 19+1 **Kieli** Englanti

Tiivistelmä

Tässä kandidaatintyössä tutkitaan satunnaisten viiveiden ja häiriöiden vaikutusta tuotannon ennustukseen eri tuotantolaitoksissa Suomessa. Työn tarkoitus on tukea tulevan tuotannon parempaa ennustamista, sekä selvittää tuotannon käyttäytymistä eri tilanteissa.

Tuotannon ennustus oli tehty yksinkertaisella mallilla, joka oletti päivittäisen tuotannon olevan vakio kuukauden sisällä lukuunottamatta niitä päiviä, jotka oli määritelty etukäteen huoltoseisokeiksi. Tutkimus toteutettiin jakamalla kuukausiennuste saman kuukauden päiville sekä laskemalla kunkin päivän toteutuneen tuotannon ja ennusteen välinen erotus. Näin oli mahdollista verrata ennusteen tarkkuutta suhteessa toteutuneeseen tuotantoon päiväkohtaisesti sekä tehdä johtopäätöksiä ennusteen tarkkuudesta laskemalla päiväkohtaisille erotuksille keskiarvo, mediaani, liukuva keskiarvo, kuvaajat luottamusväleineen sekä erittelyt päiville, joina toteutunut tuotanto oli selvästi ennustetta alempaa.

Useimpina päivinä tuotettiin enemmän kuin ennustettiin, koska ennuste on suunniteltu ottamaan huomioon tyypilliset tuotantomäärien vaihtelut. Toteutunut tuotanto oli useimpina päivinä keskiarvoaan suurempaa, kun taas keskiarvon alittavina päivinä toteutunut tuotanto oli selkeästi keskiarvoaan alhaisempaa. Päiviä, jolloin tuotantoennuste alittui, oli vähemmän, kuin päiviä, jolloin tuotantoennuste ylittyi, mutta ennusteen alittaneiden päivien toteutuneen tuotannon suuri ero ennusteesta selitti niiden suuren vaikutuksen kokonaistuotannon keskiarvoon. Tuotantoseisokkien johdosta toteutuneen tuotannon keskiarvo oli lähes sama ennusteen kanssa, kun taas toteutuneen tuotannon mediaani jäi ennustetta korkeammaksi. Merkittävää oli, että suunnittelemattomat tuotannonpysäytykset sekä pitkittyneet huoltokatkot aiheuttivat huomattavia tappioita, jolloin toteutuneen tuotannon keskiarvo oli selkeästi mediaania alempi.

Tämä tarkoittaa, että pitkillä aikaväleillä malli on hyvin luotettava, sillä toteutuneen tuotannon keskiarvo on lähes sama ennusteen antamien arvojen kanssa. Lyhyillä aikaväleillä taas ennusteen ja tuotannon välille voi syntyä huomattavia eroja. Ennustetta voisi sopeuttaa toteutuneeseen tuotantoon esimerkiksi pidentämällä huoltoseisokkien ennustettua pituutta vakiomäärällä päiviä, joka johdettaisiin päivistä, joilla huoltoseisokit pitkittyivät keskimäärin toteutuneessa tuotannossa. Huoltokuukauden ennuste voidaan puolestaan jakaa lopuille päiville, jolloin yksittäisen päivän ennuste on lähempänä toteutunutta tuotantoa.

Avainsanat Tuotanto, ennusteet, tarkkuus, tilastot, aikasarja-analyysi

Contents

1 Introduction

The purpose of this Thesis is to identify trends and outliers in monthly and daily production of 5 years of data from various production facilities of an industrial company. The primary objective was to analyze the predictability of the production in relation to the given predictions and the behavior of outlier causing situations by the model used by the company.

The raw data consisted of the past production of facilities on a daily basis for a time span of five years and was accompanied by the predictions for the production values on a monthly basis. Because the production values for each day within each month were roughly equal, it was possible to deduce the daily prediction for the production value and compare the reality and estimation of the production together, as long as planned breaks in production (maintenance, implementing investments in new technologies etc.) were taken into account. Out of these values, it was possible to calculate the estimation error within a specific time span in percentage and total tons in addition to analyzing breaks and shortages in the production.

2 Research objective

The research objective was to measure the accuracy of the given predictions of production considering the actual production values, draw conclusions based on the data and give suggestions on how to improve future predictions.

This Thesis helps understand which errors and inefficiencies in the production are so severe that they should be prioritized and dealt with before less severe ones, because there is a larger payoff in focusing on eliminating problems that cause the majority of the money loss compared to issues that do not.

Aside from these topics, it is of interest to determine statistical properties of the data to see how the predictions reflect the actual production as a whole. For instance, the mean and median of the production for a normal day or month for example above the predicted output, the predictions take into account the days when the production either lags behind the prediction or stops completely (due to e.g. equipment failure). It is also of interest to see how large the variance and standard deviation of the data are.

The production volumes of facilities are predicted before the actual production. This is done by considering how many active working days there are in a month and dividing the monthly estimate with these days. This method, however, does not take into account, for example, the time it takes to restart the facility up to full speed after a maintenance break, or how e.g. sudden machine breakdowns can affect the operation of the facilities. It is interesting to study if most of the backlog attributed to breaks is caused by a small minority of them, which cause a large portion or possibly all production to stop or whether the backlog is caused by small deficiencies in the production, which cause a minor shortage in production that go on for longer periods of time. It is also of interest to determine how the backlog days are related to each other (whether they are usually in groups of several days in a row or mostly individually in between normal days of production).

3 Methodological background

In the Thesis many basic statistical methods and concepts were used such as median, mean, moving average and graphs with 95% confidence intervals Adams and Essex (2014, p. 435-446). These were used as tools to analyze details about the behavior of the production and to assess the simple model use by the company for predicting the production. Excel and Minitab were used to analyze and edit the data. Powerpoint was used to visualize the results of the analysis.

The text books Lehtonen and Malmberg (2008); Adams and Essex (2014); Milton and Arnold (2003) provide the basis for the basic statistical methods and concepts used in this Thesis. Pace (2011) was helpful in providing backround knowledge for analyzing the data. Empie (2009) Was used as a source for basic understanding of the processes in the facilities. Taha (1992), Fitzroy et al. (1998), Krajewski and Ritzman (2002), Watson (2002), Lee and Deal (2003), Koontz et al. (1974) and Bask and Vepsalainen (1998) were used to understand the systems and overall logistics of production processes, helping to understand how unintended breaks can occur and what could cause backlog in the production generally speaking.

Time series analysis refers to the use of different methods to analyze time series data, such as production data.

The arithmetic mean (hence “mean”) is the center of weigh of the elements. Below n is the number of real-valued elements.

$$Mean = \frac{X_1 + X_2 + ... + X_n}{n}, \qquad X \in \mathbb{R}, \qquad N \in \mathbb{N}. \tag{1}$$

The median is the centermost value of the measurements in order of size. When there are two centermost values, the average of the two is taken. N is the number of elements, i.e.,

$$Median(even) = X_{\frac{n}{2}}, \qquad X \in \mathbb{R}, \qquad N \in \mathbb{N}, \tag{2}$$

$$Median(odd) = \frac{X_{\frac{n-1}{2}} + X_{\frac{n+1}{2}}}{2}, \qquad X \in \mathbb{R}, \qquad N \in \mathbb{N}. \tag{3}$$

A moving average is a method to even the random variation in a time-series. A specific number of measurements are taken around the observed measurement and the mean thereof is calculated. When this is done for every measurement in the data set, we obtain a new data set

$$MovingAverage = \frac{\sum_{i=0}^{n-1} X_i}{n}, \qquad i \in \mathbb{N}, \qquad n \in \mathbb{N}. \tag{4}$$

Confidence intervals measure the reliability of the parameter prediction, which describes the probable position of a random value from the dataset Lehtonen and Malmberg (2008); Milton and Arnold (2003). Below X is the sample mean, z the desired confidence level, s the standard deviation of the sample and n the number of elements in the sample.

$$ConfidenceInterval = X \pm z\frac{s}{\sqrt{(n)}}, \qquad X \in \mathbb{R}, \qquad s \in \mathbb{R}, \qquad n \in \mathbb{N}, \qquad 0 \leq z \leq 1. \tag{5}$$

4 Data and methods

The Thesis was done using data from four facilities, hence named Facility 1, Facility 2, Facility 3 and Facility 4. The production of these facilities was divided into subgroups e.g. by the type of raw material used. The raw data was in form of an Excel file which was modified to provide useful data for each one of the subgroups. Then, the predictions for each month were divided for every day of the month evenly. The maintenance breaks in the production were taken into account by searching for months with clearly lower predicted production than during neighbouring months and finding an equivalent gap in the true production within that month. The prediction was then set to zero for that gap and the predicted production was divided among the rest of the days of that month judged to be intended as productive. In the first iteration of analyzing the data, the daily production T and predictions H were laid side by side vertically for each subgroup and out of them was calculated the difference between them in both absolute tons and percentage difference from each other. In addition to these daily values each page received an overview for the entire addressed time span.

$$Daily\ performance\ absolute\ tons = T - H, \qquad T \in \mathbb{N}, \qquad H \in \mathbb{N} \tag{6}$$

$$Daily\ performance\ percentage = \frac{T - H}{H}, \qquad T \in \mathbb{N}, \qquad H \in \mathbb{N} \tag{7}$$

In the second iteration, a moving average of seven days was calculated for the difference between the prediction and production. When encountering a maintenance break in production, the average was calculated using values from the other side of that break as if the days making up the break did not exist. In addition to this, there were measurement categories ("Buckets") created portraying losses which occurred for at least for two days in succession and with a -20% and -50% relation of losses to the prediction for each category of production. Thus each subgrouping had two Buckets made out of them with the net loss of every continuous streak of more than one days of deficient production being classified as one element in the Bucket. The limits for meaningful deficiency being -20% and -50% respectively for both Buckets. The purpose of these Buckets was to help understand whether the difference between the prediction and production forms over time from continuous periods of minor under performance or shorter periods of major under performance.

The results in tons and percentage were taken from Excel to Minitab and were examined in various ways to analyze how well the prediction held true and in cases of unwanted breaks and days of lower production and how the these exceptional disturbances behaved e.g. whether the largest losses came during several days and how severe the loss in production capacity was daily during the loss.

In the first iteration of analysis in Matlab, the analysis of all given filtered groups for the entire given time period. Only the differences between the prediction and actual value were included in this analysis. In the second iteration, the data was divided also on a yearly basis and the -20% and -50% Buckets for accumulated losses over several days and the 7-day moving average in both tons and relative percentage for the inaccuracy of the prediction were added. Also one of the original four facilities was dropped off to streamline the analysis due to its complexities.

Finally, the results were transferred as visual presentations to PowerPoint, where the results could be observed by filtered groups within facilities and also separated for every year in addition to the total five-year period.

5 Analysis of results

In total there were 262 subgroups of data throughout all facilities not counting in Facility 4. All except 36 of the subgroups showed very similar themes with each other, which can largely be explained with small sample sizes in those 36 subgroups. Approximately half the sub groups of Buckets containing significant losses were under 10 data points for example, however they give a rough idea of what happens when the production experiences an unexpected break.

As can be expected, during longer periods the model represents the production well, though shorter periods are problematic. However, the ordinary production was usually greater than the predicted production of the facilities by a modest but steady amount, only to be dragged back by individual irregular days or mostly short series of days, when the production would reduce significantly or halt completely. Observed from the data the planned breaks also partially contributed to these irregularities by occasionally being days longer than planned and initial production afterwards getting up to speed slower than optimal. The closeness of the mean and median to the predicted production depended case by case, but usually the median of the group was far greater than the mean because there were many days, on which the actual production exceeded the predicted production slightly, whereas there were only a few days, on which the actual production was lower than the prediction, but on these days the loss in production was catastrophic. Thus, the median which takes into account the larger amount of measurements in order despite their objective size is biased towards higher values than the mean.

The nature of the production is that, by default, the production per day will remain constant and is disrupted chiefly only by surprising events that cause production to stop abruptly. It is very rare for the production to exceed its ordinary values. Exceeding the ordinary daily production is caused perhaps by underestimation of the time the facility was running during a day (for example there could have been scheduled a minor maintenance break which was supposed to last 6 hours but lasted only three). Thus, the ordinary values will always be higher than the prediction in order for the prediction itself to be accurate with the catastrophic losses that can occur on the few under performing days.

5.1 Facility 1 Results

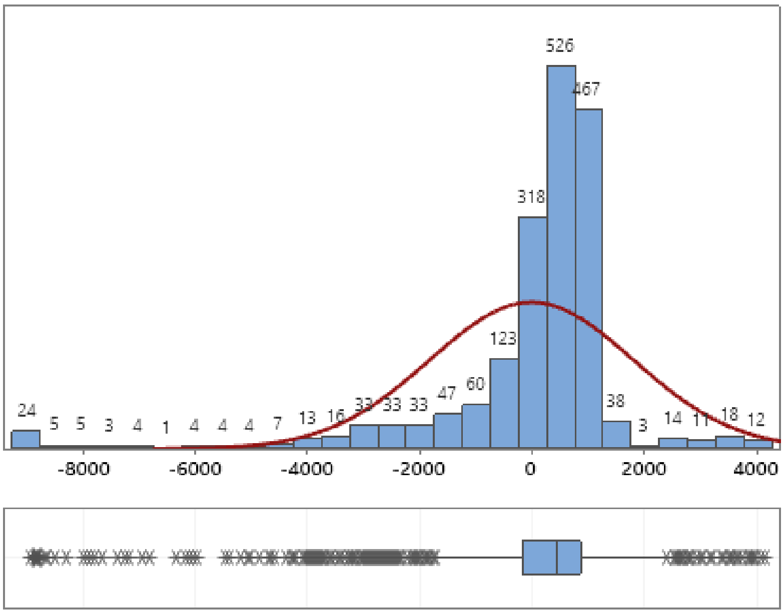

Figure 1: Boxplot of daily differences between the actual and predicted total production at Facility 1 for five years (in absolute tons).

Table 1: Output of Facility 1 in absolute tons with the original prediction subtracted.

Mean	-18,82 tons	
Standard Deviation	1810,79 tons	
Median	442,36 tons	
Confidence Interval 95% (Mean)	-101,93 tons	64,29 tons

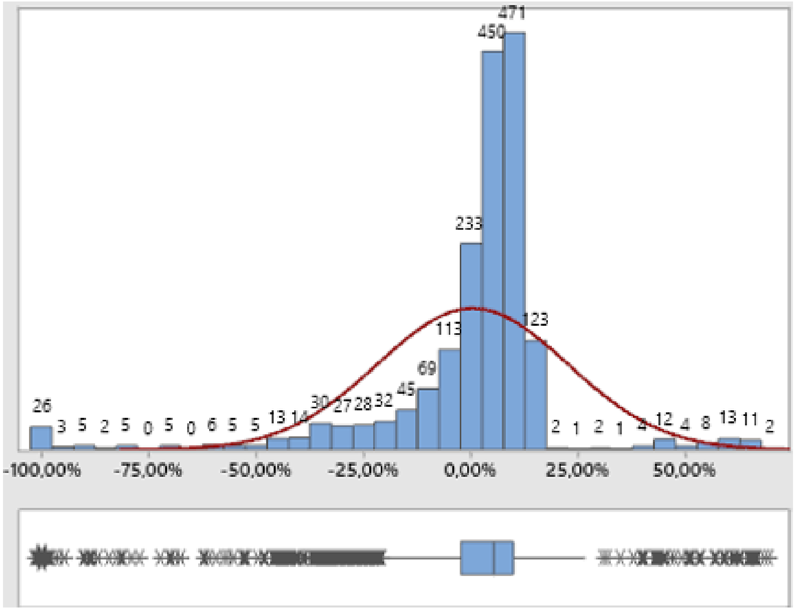

Figure 2: Boxplot of daily differences between the actual and predicted total production at Facility 1 for five years (in percentage).

Table 2: Output of Facility 1 in percentage with the original prediction subtracted.

Mean	0,322 %	
Standard Deviation	22,113 %	
Median	5,300 %	
Confidence Interval 95 % (Mean)	0,709 %	1,353 %

5.2 Facility 2 Results

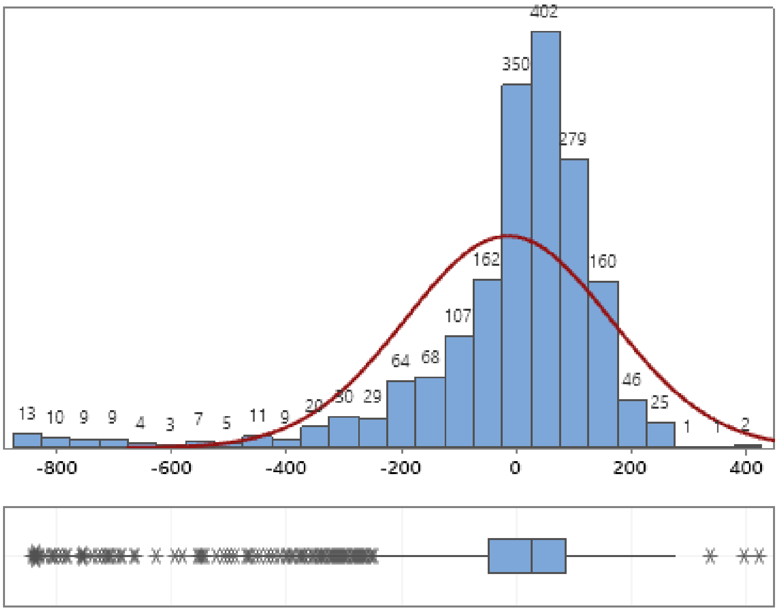

Figure 3: Boxplot of daily differences between the actual and predicted total production at Facility 2 for five years (in absolute tons).

Table 3: Output of Facility 2 in absolute tons with the original prediction subtracted.

Mean	-15,355 tons	
Standard Deviation	178,119 tons	
Median	25,215 tons	
Confidence Interval 95% (Mean)	-23,531 tons	-7,180 tons

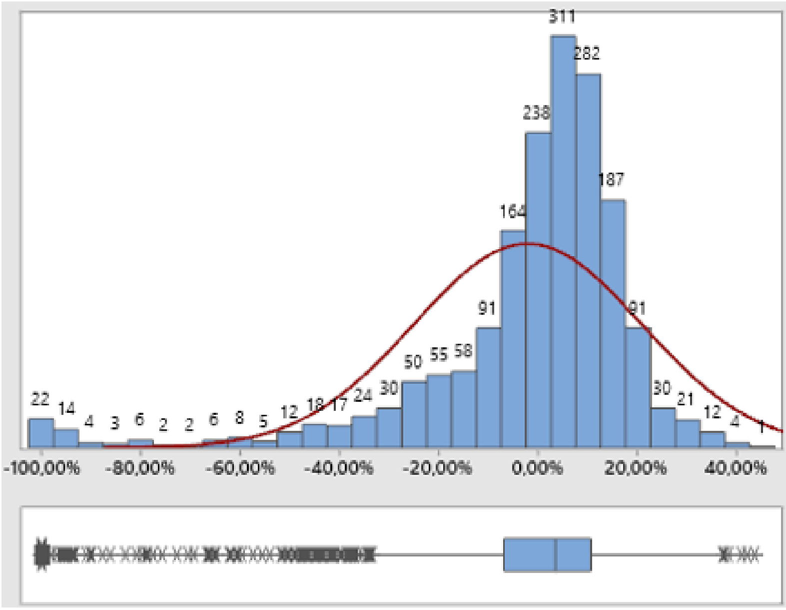

Figure 4: Boxplot of daily differences between the actual and predicted total production at Facility 2 for five years (in percentage).

Table 4: Output of Facility 2 in percentage with the original prediction subtracted.

Mean	-2,183 %	
Standard Deviation	22,877 %	
Median	3,385 %	
Confidence Interval 95% (Mean)	-3,250 %	-1,116 %

5.3 Facility 2 Moving Averages

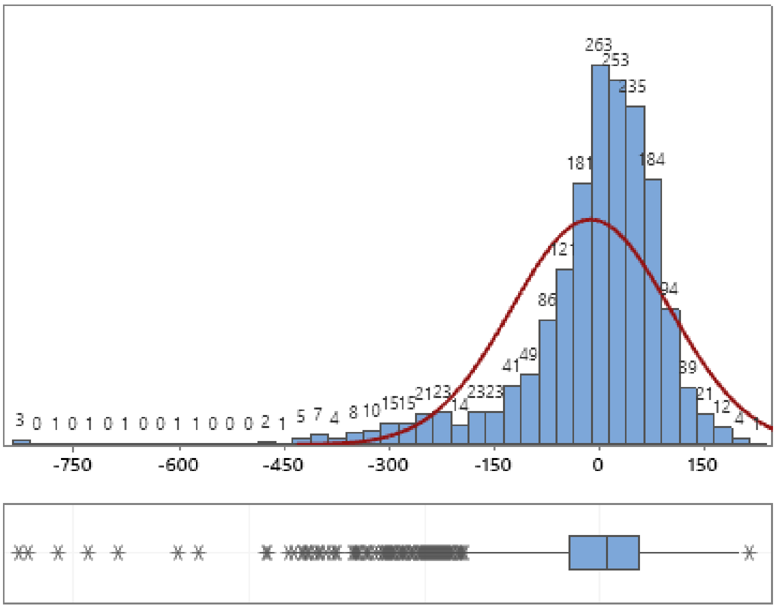

Figure 5: Boxplot of moving averages at Facility 2 for five years (in absolute tons).

Table 5: Moving averages of Facility 2 (in absolute tons).

Mean	-13,498 tons	
Standard Deviation	112,650 tons	
Median	9,352 tons	
Confidence Interval 95% (Mean)	-18,760 tons	-8,236 tons

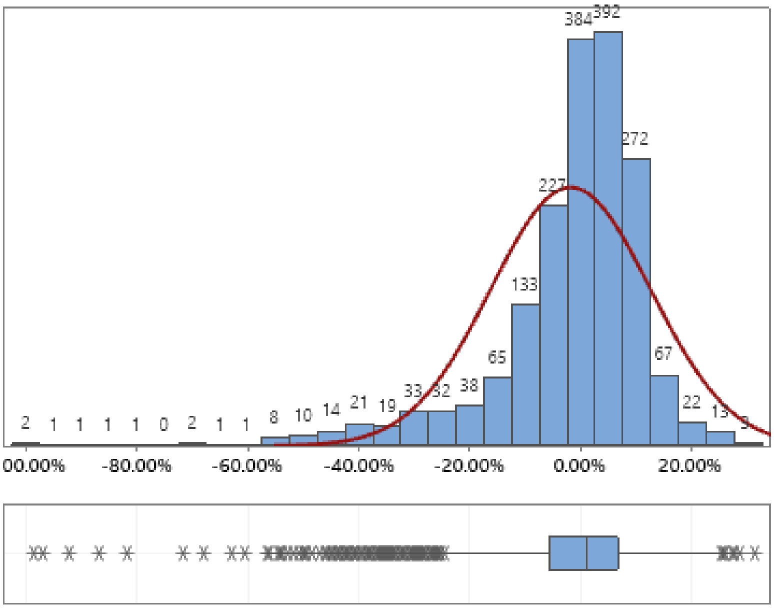

Figure 6: Boxplot of moving averages at Facility 2 for five years (in percentage).

Table 6: Moving averages of Facility 2 (in percentage).

Mean	-1,889 %	
Standard Deviation	14,387 %	
Median	1,066 %	
Confidence Interval 95% (Mean)	-2,561 %	-1,217 %

5.4 Facility 2 Buckets

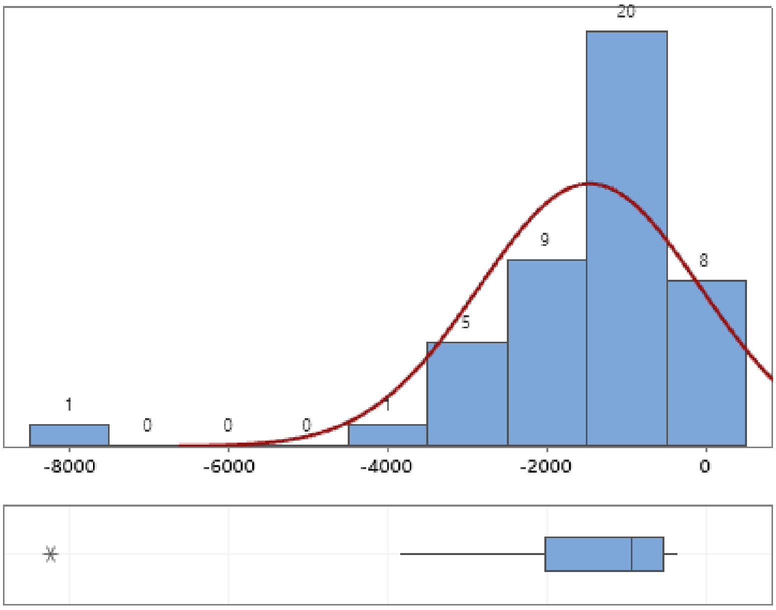

Figure 7: Boxplot of the Bucket of -20% underperforming periods (sorted by the amount of loss).

Table 7: The sample size of the Bucket of -20% underperforming periods is small. The worst performing periods are a major reason for drops in overall production.

Mean	-1468,6 tons	
Standard Deviation	1387,7 tons	
Median	-934,9 tons	
Confidence Interval 95% (Mean)	-1890,5 tons	-1046,7 tons

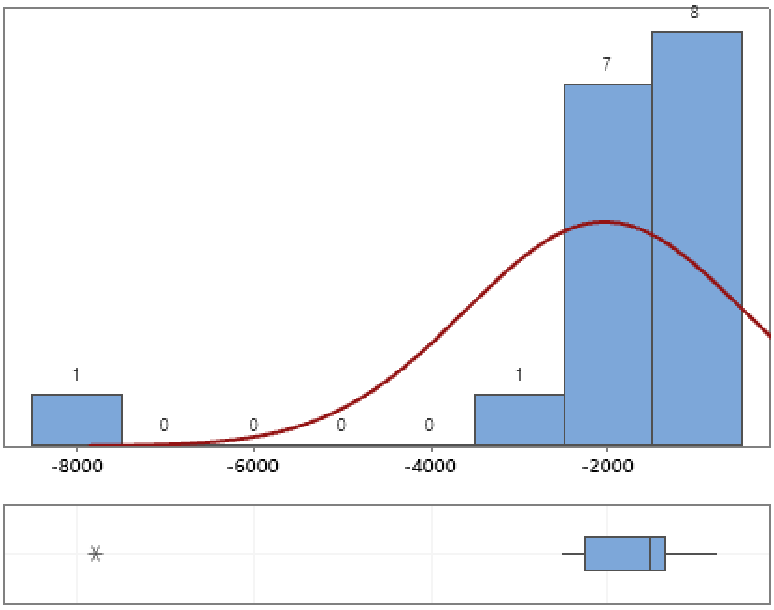

Figure 8: Boxplot of the Bucket of -50% underperforming periods (sorted by the amount of loss).

Table 8: The sample size of the Bucket of -50% underperforming periods is small. The worst performing periods are a major reason for drops in overall production.

Mean	-2044,2 tons	
Standard Deviation	1566,3 tons	
Median	-1517,7 tons	
Confidence Interval 95% (Mean)	-2849,5 tons	-1238,8 tons

6 Conclusion

Overall, the results in the facilities and subgroups have many similarities and the differences can be explained through peculiarities within each group. Many of the groupings especially for the Buckets of days with continuous inconsistency between prediction and production suffered from small sample sizes and thus were necessarily not completely representative of the production. However, they give a picture of how shortages and halts in the production behave.

The results show that on most days, the predicted production is well below the actual production of a normal day of production, due to the days, when predictions were not met, dragging both the mean and median of the production down with the mean being more severely affected.

The original simple prediction where the monthly prediction was divided by days of production was reasonably accurate, but could be fine-tuned around the planned maintenance breaks of the operations, as production efficiency seems to lag often for a few days after the continuation of production. Also, one could analyse what the daily normal production (or monthly production, assuming everything works perfectly) would be separately compared to the current model, where the model assumes the stops in production to occur, thus giving an underwhelming daily prediction for the production. Seeing this difference could help in e.g. making decisions concerning investments into the reliability of the machinery and process in general.

References

R. Adams and C. Essex. *Calculus A Complete Course 8th edition.* Pearson, Toronto, 2014.

A. Bask and P. Vepsalainen. *Opening Markets for Logistics.* The Finnish Association of Logistics, Helsinki, 1998.

H. Empie. *Fundamentals of the Craft Recovery Process.* Tappi Press, Norcross, GA, 2009.

F. Fitzroy, Z. Acs, and D. Gerlowski. *Management of Economics and Organization.* Prentice Hall Europe, Hertfordshire, 1998.

H. Koontz, C. O'Donnell, and H. Weihrich. *Essentials of Management.* McGraw-Hill, United States, 1974.

L. Krajewski and L. Ritzman. *Operations Management Strategy and Analysis.* Pearson Education, New Jersey, 2002.

G. Lee and T. Deal. *Reframing Organizations.* John Wiley Sons, San Fransisco, 2003.

T. Lehtonen and J. Malmberg. *Tilastotieteen alkeita.* Hakapaino Oy, Helsinki, 2008.

J. Milton and J. Arnold. *Introduction to Probability and Statistics*. McGraw-Hill, New York, 2003.

L. Pace. *Statistical Analysis Using Excel 2007*. Pearson Education, Upper Saddle River, 2011.

H. Taha. *Operation Research An Introduction 5th edition*. Macmillan Publishing Company, New York, 1992.

T. Watson. *Organising and Managing Work*. Pearson Education, Harlow, 2002.