mathematics

MDPI

Article

Use of Statistical Process Control for Coking Time Monitoring

Marta Benková, Dagmar Bednárová, Gabriela Bogdanovská * and Marcela Pavlíčková

Institute of Control and Informatization of Production Processes, Faculty of Mining, Ecology, Process Control and Geotechnologies, Technical University of Košice, Němcovej 3, 042 00 Košice, Slovakia; marta.benkova@tuke.sk (M.B.); dagmar.bednarova@tuke.sk (D.B.); marcela.pavlickova@tuke.sk (M.P.)

* Correspondence: gabriela.bogdanovska@tuke.sk; Tel.: +421-602-5170

Abstract: Technical and technological developments in recent decades have stimulated the rapid development of methods and tools in the field of statistical process quality control, which also includes control charts. The principle of control charts defined by Dr. W. Shewhart has been known for more than 100 years. Since then, they have been used in many industries to monitor and control processes. This paper aims to assess the possibilities of use and the selection of the most suitable type of control chart for monitoring the quality of a process depending on its nature. This tool should help operators in monitoring coking time, which is one of the important control variables affecting the quality of coke production. The autoregressive nature of the variable being monitored was considered when selecting a suitable control chart from the group of options considered. In addition to the three traditional types of control charts (Shewhart's, CUSUM, and EWMA), which were applied to the residuals of individual values of different types of ARIMA models, various statistical tests, and plots, a dynamic EWMA control chart was also used. Its advantage over traditional control charts applied to residuals is that it works with directly measured coking time data. This chart is intended to serve as a method to monitor the process. Its role is only to alert the process operator to the occurrence of problems with the length of the coking time.

Keywords: statistical process control; control chart; autocorrelated process; ARIMA model; statistical test; coking time

MSC: 62P30

check for updates

Citation: Benková, M.; Bednárová, D.; Bogdanovská, G.; Pavlíčková, M. Use of Statistical Process Control for Coking Time Monitoring. *Mathematics* **2023**, *11*, 3444. https://doi.org/10.3390/math11163444

Academic Editor: Don Wardell

Received: 30 June 2023
Revised: 31 July 2023
Accepted: 4 August 2023
Published: 8 August 2023

1. Introduction

If a product is to meet customer needs and expectations, it should be created using a stable process, i.e., a process that is capable of operating with little variability around the target product characteristics. Therefore, statistical process quality control has been used in industrial practice for almost 100 years. Montgomery calls it one of the greatest technological achievements of the 20th century, given that it is easy to use, it has a significant impact, and can be applied to a variety of processes [1]. Its basic method is the use of control charts, which are based on a prevention strategy.

There is some variability in each process. The study of the various causes of this variability enables us to understand the process better and to manage it in such a way as to ensure that its outputs conform to the specified requirements. Maintaining a process in a stable state, i.e., a state in which only a stable system of random or common causes operates, is called a state of statistical control. The use of control charts allows the process to be kept in a state of statistical control and thus prevents the occurrence of unusable non-deterministic outputs of the process. It also makes it possible to identify, investigate, and reduce the impact of systematic or attributable causes of variability that can negatively affect the quality of process outputs.

A control chart is indeed an effective statistical control tool if a suitable one for the process under study is selected from the many control charts that exist today and its parameters are appropriately set.

1.1. Traditional Control Charts

The first control charts for variables and control charts for attributes were designed and introduced by Shewhart. They were developed for one regulated quality characteristic [2]. Since they are control charts without memory, they are suitable for identifying large shifts in the mean, higher than three standard deviations, in the process.

The cumulative sum control chart (CUSUM) was proposed by Page [3]. Several authors, e.g., Lucas [4], Hawkins [5] confirmed its effectiveness for identifying small shifts in the process mean due to uniform and unlimited memory.

Roberts [6] introduced the exponentially weighted moving average control chart (EWMA). Crowder [7,8] and Lucas and Saccucci [9] evaluated it as a suitable alternative to the Shewhart control chart for the identification of small shifts in the mean, less than 0.5–2.0 standard deviations. It is a control chart with uneven and unlimited memory. The memory is controlled by the size of the smoothing parameter λ, where $0 < \lambda < 1$. Small λ values are useful for identifying small shifts in the process, while large values of λ better identify larger changes. Zhang et al. proposed the EWMA control chart, which is more robust in avoiding estimation errors or standard deviation changes [10].

The basic assumptions for using the above three traditional control charts for variables are normality, constant mean, homoscedasticity, and independence of the data of the regulated quality characteristics. Failure to comply with these assumptions can negatively affect their performance [11–13].

The sample size and the length of the sampling frequency also play an important role in the effective use of control charts. Larger samples are more effective for identifying smaller process shifts. However, current practice prefers rather small subgroups, which are implemented more frequently, aided by automated measurement systems [14].

1.2. The Phases of Using Control Charts

The application of the control chart involves two phases—the retrospective phase (Phase I) and the monitoring phase (Phase II). Each phase has a different content and objective.

In Phase I, measured data from about 20–25 samples are used to retrospectively analyse the process behaviour and to calculate regulatory limits for the next phase. The use of a control chart is a repeated test of the hypothesis that the process under study is under statistical control. Thus, it is associated with the occurrence of a Type I error, called the risk of a useless signal (α), and with the occurrence of a Type II error, called the risk of a missing signal (β) [11]. Very often, control limits at a distance of three sigma from the central line of the control chart (three-sigma control limits) are used. According to Shewhart, mathematical theory, empirical evidence, and practical experience provide evidence that it is the three-sigma limits that minimize the chance of a Type I or II error [2]. Costa and Fichera state in their study that originally the design of control charts only considered statistical criteria [15]. The aim was to reduce the occurrence of Type I and Type II errors. From the 1980s onwards, the economic aspect came to the fore. Nowadays, the economic–statistical approach is preferred. In Phase I, which may involve several cycles, the systematic causes of variability are gradually removed, and the control limits are revised by means of a control chart.

Phase II begins when the process is already under statistical control. The calculated control limits are used to monitor it using the newly measured data for each sample. In this phase, it is useful to focus attention on monitoring patterns in control charts [16]. These clusters of out-of-control points can be visually identified in the control chart and their cause in the controlled process can be removed in time. For these activities, process operators often have an out-of-control action plan [14]. Other groupings of points such as cyclic patterns, mixtures, process level shifts, or trends can also be found in control charts. Cyclic patterns are related to systematic changes in the environment, e.g., temperature, pressure fluctuations, voltage, or operator changes. Kalteh and Babouei presented a new method for recognizing nine types of patterns in control charts based on the use of shape and statistical features and an optimized fuzzy system [17]. For abnormality detection, Fuqua and Razzaghi described a cost-sensitive classification scheme in the framework of a deep convolutional neural network for the recognition of patterns in a control chart. This approach is very suitable for intelligent manufacturing systems that work with large datasets. Recognition algorithms for control chart patterns can make it easier to monitor product quality [18]. Garcia et al. identified six basic recognition approaches to control chart patterns [19]. In their review, they mapped the recognition of 21 different control chart patterns in 41 publications to date, 11% of which included autocorrelated data. If the existence of a trend visible in a control chart is implied by the nature of the process itself, it is appropriate to use a regression control chart as reported by Shu et al. [20], Hayati [21], and described in [22]. The faster the action of the systematic cause of variability is identified in Phase II, the more effective the control chart is. The ability of the control chart to quickly detect the existence of systematic causes of variability can provide clues to more quickly and efficiently identify the source of the quality problem [23]. Therefore, as Cuentas et al. state, current trends in statistical process control include the use of machine learning algorithms [24].

Montgomery recommends the use of Shewhart control charts in Phase I [14]. Their advantages are simple calculations and high efficiency in detecting large and permanent shifts in process parameters, i.e., 2σ. In Phase II, small to moderate shifts in process parameters are often encountered. Shewhart control charts are then less effective, which can lead to an increased false-alarm rate. Much more effective at this stage are the CUSUM [25,26] and EWMA control charts [7–9,27]. Shewhart control diagrams always consider only the results of the last sample. The CUSUM and EWMA control charts consider the control results of the last sample, but also those of the preceding samples. CUSUM considers the results of the previous subgroups with equal weight. EWMA allows the weights of the previous subgroups to be set unequally using a smoothing parameter. Therefore, CUSUM and EWMA control charts are more sensitive to small process variations than the Shewhart control chart.

1.3. The Problems Caused by Autocorrelation of Observed Data

The Shewhart control chart, the CUSUM control chart, as well as the EWMA control chart are based on the assumption that the processed data are statistically independent. As several authors state in their papers, e.g., [28,29], this assumption is not valid in all industries, e.g., for continuous processes in the metallurgical, chemical, and food industries, but also for highly automated discrete manufacturing processes in the mechanical or electrical engineering industries. According to [27], processes with dependent observations can be divided into two basic types, which are

- Stationary autocorrelated processes, which are in a state of so-called statistical equilibrium. The basic behaviour of these processes does not change over time, neither do the mean values and deviations.
- Non-stationary autocorrelated processes, which are not in a state of statistical equilibrium. Their means and variances change, and it is often possible to observe the occurrence of a trend.

As stated by [14,28,29], verifying the assumption of the independence of observations is very important. Traditional control charts can give misleading results due to the large number of false signals. Many authors pointed out the influence of correlation. Patel and Divecha [30] state that the typical effect of autocorrelation is to reduce the average run length, leading to a higher false-alarm rate than in the case of an independent process (an explanation of the concept of average run length is included in Section 1.5). Cheng et al. state that the presence of autocorrelation in the measured data affects the statistical performance of control charts. It creates control limits that are much more stringent than would be desirable. This causes an increase in the number of false signals and a decrease in the reliability of the control chart application [31].

1.4. The Approaches to Solving Problems with Autocorrelated Data

Many authors use different approaches to solve problems in autocorrelated processes using industrial applications of control charts.

Zhang applied the ARMA time series model to stationary autocorrelated processes, denoted as an ARMAST chart, and an EWMAST chart [32]. In the next study, Zhang compared EWMAST control charts, CUSUM control charts for residuals, and EWMA control charts for residuals, using the average run length [33].

Woodal and Montgomery tracked the autocorrelated process level using a time series model or using a control chart based on one-step-ahead forecast errors [34]. Shu et al. recommend the use of the triggered cuscore chart or the general likelihood-ratio test instead of the traditional control charts for residuals [35]. The adaptive control chart proposed by Capizzi and Masarotto weights past observations of the process under study using an appropriate function of the current "error". The proposed control chart is effective in detecting shifts of different magnitudes, but the monitoring scheme is quite complicated [36]. Testik reports that common approaches for autocorrelated observations are to explain the dynamics of the process by an appropriate time series model and to monitor the predicted residuals to be independent, or to monitor the autocorrelated observations using a control chart with adjusted control limits [37]. Castagliola and Tsung investigated the impact of skewness data from an autocorrelated process on traditional control charts and provided a scaled weighted variance approach to improve their performance by using a scaled weighted variance method [38].

Traditional control charts use samples of equal size measured at the same control interval. Zou et al. used an adaptive control chart with a variable sample size and fixed control interval for an autocorrelated process [39].

Patel and Divecha proposed a modified EWMA control chart. By combining the Shewhart control chart and the EWMA control chart, they were able to identify both small and large shifts. Some industrial processes with high levels of first-order autocorrelation require this approach. Its basic idea is to adjust the weight of past observations, past changes, current observations, and current change [30]. Khan et al. presented their version of the modified EWMA control chart [40]. It is based on the statistic published by Patel and Divecha and can be applied to autocorrelated data.

Evaluating clean coal quality by monitoring the clean coal ash content using the Shewhart control chart is discussed by Fu et al. in their study. Due to the occurrence of a significant autocorrelation of the monitored parameter, ten types of data collection procedure were tested and compared in order to reduce the influence of autocorrelation on the determination of the standard deviation of the process [41].

Osei-Aning et al. selected optimal parameters for the EWMA and CUSUM control charts using an exhaustive search procedure that is suitable for monitoring autocorrelated data. They determine the parameters that produce the smallest extra quadratic loss for each autocorrelation coefficient. The paper provides optimal parameters that can be used to increase the overall efficiency of control charts [42].

Li et al. present three different approaches. The first approach is to reduce the frequency of data collection. The second approach is based on re-evaluating the actual variance in the process and adjusting the control limits. The last one recommends the use of residual charts. The authors described an improved hidden Markov model based on a residual chart for monitoring the autocorrelated process. The conducted experiment shows that this model outperforms both the conventional hidden Markov model and autoregressive models in diagnosing quality shifts, thus reducing the cost of missing signals [43].

Yao et al. [44] identified two main approaches for monitoring autocorrelated data. The first approach is to use control charts for residuals of the time series model. However, the authors point out that this approach introduces problems with the interpretability and effectiveness of control charts and the lack of detection of assignable causes. The second approach uses data from the autocorrelated process directly, and also uses time series models and takes into account the structure of variance and covariance when calculating the control limits.

Reynolds et al. recommend investigating whether autocorrelation is a special cause or whether it is part of common-cause variability. In the first case, the autocorrelation should be removed and in the second case the autocorrelation should be accounted for in the control chart. These authors distinguish two approaches to modelling autocorrelated processes. The first involves the use of traditional control charts with modified control limits. The second approach uses the residuals of time series models or prediction errors in the regulation diagrams [45]. The level of autocorrelation should also be taken into account. According to Yashchin [46], the first approach works better at low to medium levels of autocorrelation, while the second approach will work better at high levels of autocorrelation.

Based on the above sources from the literature, approaches to statistical process control for autocorrelated processes can be summarized as follows:

- The extension of the control interval;
- The use of the traditional control charts with modified control limits;
- The use of the traditional control charts, or slightly modified versions of them, for residuals of time series models that are not autocorrelated;
- The use of the non-parametric control charts.

These options are analysed in more detail in the following subsections.

1.4.1. The Extension of the Control Interval

Sampling at a lower frequency, i.e., extending the control interval, is the simplest solution to remove autocorrelation in data. The appropriate length of the control interval for a particular process should be verified by the autocorrelation function. However, this procedure leads to a loss of information about the behaviour of the process. If, for example, only every ninth observation is used, this means a loss of 90 percent of the data. It may reduce the performance of the control chart used by taking longer to detect the actual process drift than with a shorter interval [14,47,48].

Other options present solutions that are more complex. For example, Franco et al. proposed a new mixed sampling strategy for the Shewhart chart. The mean value of the sample at each control time is calculated by combining the measurements of the quality characteristic from two consecutive samples taken hours apart [49]. Grimshaw proposes exhaustive systematic sampling that is similar to Bayesian thinning. Practical guidance is offered for selecting a systematic sampling interval large enough to be approximately unbiased and not too large to increase the variance [50].

1.4.2. The Use of the Traditional Control Charts with Modified Control Limits

This approach uses the computation of control limits in such way that they respect the autocorrelation structure of the data by using their distance from the central line, while having the ability to provide an out-of-control signal fast enough. This requires a correct estimation of the real standard deviation of the autocorrelated process, for example [51]. Osei-Aning et al. present modified Mixed EWMA-CUSUM and Mixed CUSUM-EWMA control charts. The authors modified the control limits of traditional graphs to better monitor autocorrelated data [42]. Garza-Venegaz et al. developed a bootstrapping technique to adjust the control limits [52].

1.4.3. The Use of the Traditional Control Charts for Residuals of Time Series Models

The application of traditional control charts to the residuals of a time series model computed according to the Box–Jenkins methodology appears in the work of many authors. Individual authors using this approach differ in the type of time series model used.

Alwan and Roberts recommend the use of a common-cause chart, e.g., a chart of forecasted values that are determined by fitting the correlated process with an ARIMA model, and a special-cause chart, e.g., a traditional control chart of the residuals or one-step-ahead prediction errors [53]. Wardell et al. followed up on their work [54,55]. They showed that a special-cause diagram is more likely to detect process shifts quickly than traditional control charts.

Lu et al. applied the EWMA control chart to the process residuals obtained as an autoregressive model of the ARIMA (1,0,0) process and evaluated it using the integral equation method. They found that for high levels of autocorrelation and large displacements, the EWMA control chart applied to the residuals is faster than the EWMA control chart applied to the original observations [56]. Jiang et al. described the autoregressive moving average control chart and approach to monitoring autocorrelated data [57]. Lu et al. confirmed in their study the appropriateness of using CUSUM and EWMA control charts for monitoring the observations from the process as an autoregressive process [58]. Tasdemir presents the effect of autocorrelation on process control charts to monitor two quality characteristics of fine coal produced in a coal washer for a power plant. He used an ARIMA (1,0,1) model for moisture content and an ARIMA (0,1,2) model for ash content. He found the above time series models to be the best models to remove autocorrelation [59]. Magaji et al. described the application of the EWMA control chart to the residuals of the ARIMA (2,0,0) model obtained from autocorrelated chemical process viscosity data [60]. Li et al. used and compared with each other two types of exponentially weighted moving averages of autoregressive model residual (EWMA-R) control charts to monitor the autocorrelated process of order p [48]. Costa and Fichera describe the design of an ARMA control chart. To select the optimal model parameters, they developed a modified self-adaptive differential evolution algorithm [61]. Li et al. pointed out that the use of ARIMA models requires quite a lot of experience, which sometimes causes inconvenience in implementation. They proposed hidden Markov models and showed that these models are more stable than first-order autoregressive models in the case of both positive and negative autocorrelations of observations [62]. Phanthuna and Areepong used a modified EWMA control chart for an integrated moving average model and a fractional integrated moving average model applied to datasets of natural gas and crude oil prices [63].

Mastrangelo and Montgomery presented a method based on an exponentially weighted moving average that uses variable control limits. Their approach assumes that the autocorrelated process can be modelled using an ARIMA (0,1,1) model. A control chart for individual values is used for one-step prediction errors, i.e., to the residuals of the EWMA model. The proposed control chart is only suitable for processes with a positive autocorrelation of the non-constant mean with slow drift; a rapid change in the mean value of the process leads to an exceedance of the control limits [64,65].

One of the simplest non-parametric control charts, a Shewhart sign control chart for tracking the position of a process, has been described, e.g., by Armin et al. [66]. Chakraboti et al. published a review of non-parametric control charts, highlighting their various advantages [67]. Chakraboti and Graham, and Chakraboti and Eryilmaz, considered the main advantage of non-parametric control charts to be their general flexibility. Their application does not require meeting the assumption of any specific probability distribution for the measured data [68,69]. They assume only a continuous distribution. Therefore, the term "distribution-free" is more appropriate than "non-parametric" for this type of control chart. They are less sensitive to the random occurrence of outliers and are often simpler than traditional control charts. Figueiredo and Gomes compared robust control charts for mean and standard deviation tracking using Monte Carlo simulations in terms of their robustness and performance [70]. Bakir presents an extensive review of the literature related to non-parametric control charts, spanning a period of almost 80 years [71]. Kountras and Triantafyllou showed a general procedure for constructing non-parametric regression charts. They use specific order statistics to determine appropriate control limits. To decide whether a process was under control, they used order-based statistics as well. They introduced three new non-parametric control charts based on minimum and maximum rank statistics [72]. Smajdorová and Noskievičová stated that non-parametric control charts are suitable for monitoring intelligent manufacturing processes with complex structures based on a large amount of data. Their study contains an algorithm for selecting a suitable non-parametric control chart according to type of the distribution, skewness, and kurtosis of the process data [73].

1.5. The Evaluate the Performance of Control Charts

An important part of using control charts is to evaluate their performance, i.e., their ability to detect changes of a certain magnitude in the monitored process parameter as quickly as possible. It helps to select the most appropriate control chart for a particular process. Different performance indicators are used for this evaluation [1,14,74–78]. Most of them are based on the run length, i.e., on the number of points that are recorded in the control chart, before the out-of-control condition is identified. Average run length (*ARL)* is the most used performance indicator. This indicator is also useful for evaluating different subgroup sizes and control interval lengths [14]. For a process that is under control, *ARL(0)* is used; for a process that is out of control, *ARL(1)* is used. *ARL(δ)* is also used, where the value δ represents the specific magnitude of the shift in the mean value. When designing a control chart, it is generally required that the value of *ARL(0)* be as large as possible and the value of *ARL(δ),* conversely, be as small as possible [74].

Low et al. suggested using the median run length (*MRL)* value as a more robust performance indicator [75]. Chin and Khoo state that *MRL* provides a more meaningful interpretation of the graph results for under-control and out-of-control processes. Quality practitioners easily understand *MRL* because it provides the probability of a signal at a certain number of samples [76]. Jones et al. [77] or Antzoulakos and Rakitzis [78] consider the standard deviation of run length (*SDRL)* as a complementary indicator to assess the performance of control charts.

Graphical methods are also used to evaluate the performance of control charts. The *ARL* curve shows the dependence of the average run length on the magnitude of the process parameter shift δ expressed as the number of standard deviations. The operating characteristic graph shows the dependence of the risk β on the magnitude of the process parameter shift δ [74].

1.6. The Purpose of the Work

The study aims to assess the possibilities of the active use of SPC tools for monitoring a technological process. The use of traditional control charts is based on several basic assumptions. One of them is that the data characterizing the process are statistically independent. In practice, however, there are many processes that are characterized by dependent, autocorrelated data. In such cases, several solution options can be taken into consideration, which have already been described in the previous section. Our main purpose is to assess and compare their possible applications and to select a suitable tool for application in a particular process.

2. Materials and Methods

The first subsection briefly describes the coking process and one of its important control variables, which is the coking time. The datasets of the coking time used for the calculations are described in the second subsection. It is this variable that is further monitored by means of several types of control charts, which are explained in more detail in the next subsections.

2.1. The Coking Process

Metallurgical coke production takes place in classical multi-chamber coke oven batteries (Figure 1a), via high-temperature carbonisation of the coal charge at a temperature of 1000 °C in the absence of air [79]. The aim of the conversion of coal into metallurgical coke (Figure 1b) is to produce a material rich in carbon and at the same time free of the chemical impurities present in the coal in the coal charge. The operation of the coke battery should be steady with the optimum amount of heat consumed and the desired quality of coke achieved. The quality and preparation of raw hard coal, the heating conditions, and several technological control variables must be in balance. Even a small imbalance can lead to a loss of stability of the whole coking process. The fulfilment of these requirements requires the establishment of a technological regime for the dependence on coking time and coking temperature [80].

(**a**) (**b**)

Figure 1. (**a**) Illustration of a multi-chamber coke oven battery; (**b**) illustration of metallurgical coke.

The coking time is the time taken for the carbonization of the coal charge in the chamber in hours, from the occupation of the chamber by the charge to the obtaining of high-temperature coke in the whole cross-section of the chamber. The complete coke-making process takes approximately 20–24 h. Its length is influenced by

- The width of the coking chamber (the wider the chamber, the longer the coking time);
- The temperature of the heating walls (the higher the temperature in the heating channels of the walls, the shorter the coking time);
- The quality and thickness of the heating walls (i.e., the distance between the channel and the chamber depends on the type and quality of the dinas used, with a 10 mm change in thickness caused by changing the coking time by 0.6–0.8 h);
- The quality of the coal charge (moisture content, grain composition, and density of the charge have the greatest influence on the heat transfer through the charge, and increasing these increases the coking time;
- Coking properties also change the coking time, a change of 5% in the proportion of coking coals requires a change in coking time of 1.0–1.5 h [81].

In this paper, only coking time data are processed. Other important technological control variables of coke oven battery operation, such as heating, exhaust, control of source and equipment operators, etc., are not dealt with.

2.2. Description of the Data Used

The automated control system for the operating machines of the coke oven battery that contains thirty chambers divided into three equal blocks (Figure 1b) is made up of a set of programmable logic controller (PLC) machines and a computer. The main task of this system is to control the operation of the operating machines (extruders, guide cars, filling cars, and coke extinguishing technology sets), but the real-time recording of operational activities is an equally relevant task. The most important data measured and collected include coking time, as one of the quality characteristics of the technological process of coke production. The PLC machine, which controls the time sequence of the work of the machines, records the date and time of occupation and the date and time of coke extrusion from the chamber. The difference between these times is the coking time, which then enters statistical processing.

This study used coking time data for three selected days in two consecutive months (January and February) of the same year. The January datasets were used for baseline validation of the assumptions about the investigated coking time and for the calculation of the parameters of the control charts, i.e., for the retrospective analysis—Phase I. The February datasets were used for process monitoring when control charts with calculated control limits were used for the monitoring phase—Phase II. Each dataset contains approximately one hundred measured values of coking time.

2.3. Traditional Control Charts Calculations for Individual Measurements

The section describes the several control charts that later were used for the practical calculations in Section 3, Results and Discussion. The charts for individual measurements were used because the measured data of the variable under study, which is the coking time, refer to different chambers of the coke battery.

2.3.1. Shewhart Control Chart for Individual Measurements

The upper control limit (UCL), centre line (CL), and lower control limit (LCL) for the Shewhart control chart for an individual can be calculated using the following formulas:

$$\begin{aligned} UCL &= CL + L \cdot \frac{\overline{MR}}{d_2} \\ CL &= \mu_0 \\ UCL &= CL - L \cdot \frac{\overline{MR}}{d_2} \end{aligned} \tag{1}$$

where μ_0 is the target mean value. $\overline{MR}$ is the average moving range calculated using Formula (2), parameter L for the control limits is equal to 3, and d_2 is the constant for the construction control chart for n = 2 [16].

$$\overline{MR} = \frac{1}{n}\sum_{i=1}^{n} MR_i = \frac{1}{n}\sum_{i=1}^{n}|x_i - x_{i-1}| \tag{2}$$

2.3.2. Tabular CUSUM Control Chart for Individual Measurements

For tabular CUSUM control charts, the one-side upper and one-side lower CUSUM statistics are calculated according to the following formulas:

$$\begin{aligned} C_i^+ &= \max\left[0, x_i - (\mu_0 + K) + C_{i-1}^+\right] \\ C_i^- &= \max\left[0, (\mu_0 - K) - x_i + C_{i-1}^-\right] \end{aligned} ; C_0^+ = C_0^- = 0 \tag{3}$$

where the reference value K is a constant set by the user [5]. Very often, it is calculated using the half-way point between the target mean value μ_0 and the out-of-control value $\mu_1 = \mu_0 \pm \delta \cdot \sigma$, which need to be detected quickly, i.e., according to the formula

$$K = \frac{|\delta \cdot \sigma|}{2} = \frac{|\mu_1 - \mu_0|}{2}, \tag{4}$$

where μ_1 is the mean out-of-control value. C_i^+ and C_i^- represent deviations from μ_0 that are greater than K. Both start at zero and reset to zero again after becoming negative values [82]. The setting of the decision interval parameters H and reference value K for a particular process affects the effective use of the CUSUM control chart. Hawkins lists the ranges of k $(K = k \cdot \sigma)$ and h $(H = h \cdot \sigma)$, where σ is the standard deviation of the variable values that are suitable for achieving an average run length equal to 370 [83].

2.3.3. EWMA Control Chart for Individual Measurements

The EWMA statistic is calculated as the exponentially weighted moving average of the following formula:

$$z_i = \lambda \cdot x_i + (1 - \lambda) \cdot z_{i-1}, \text{ with } z_0 = \mu_0; \tag{5}$$

where λ is a constant chosen by the user from 0 to 1 [14]. The starting value of EWMA for the first step $i = 1$, is the process target μ_0. Because the used weights decline geometrically, EWMA is called a geometric moving average. The control limits are calculated according to the following formulas:

$$\begin{aligned} &UCL_i = CL + L \cdot \sigma_i, \\ &CL = \mu_0, \\ &LCL_i = \mu_0 - L \cdot \sigma_i; \end{aligned} \tag{6}$$

these are narrowed limits where standard deviation σ_i is calculated as

$$\sigma_i = \sigma \cdot \sqrt{\frac{\lambda}{2-\lambda} \cdot \left[1 - (1-\lambda)^{2 \cdot i}\right]}. \tag{7}$$

The control limits are initially narrower, gradually widening, and after a few periods, they approach the steady state. This is due to the part of the formula in square brackets approaching the unit. The used values of the parameters L and λ are important for the efficiency of the EWMA control chart. Small values of λ are useful for identifying small shifts, while large values of λ better identify large changes. The values $0.05 \leq \lambda \leq 0.25$ are often used. Montgomery warns about the inertia effect when low values of λ are used, which can reduce the effectiveness of EWMA control chart. When a shift occurs on the opposite side of the centre line, it could take several periods for the EWMA values to respond to the shift, since a small value of λ does not weigh the new data very heavily [14]. A convenient way to find the optimal value for the investigated process is to use the nomograms created by Crowder [7] for $ARL(0)$ equal to 50, 100, 250, 370, 750, 1000, 1500, and 2000. Very often, $ARL(0) = 370$ is used because it corresponds to α = 0.0027. Generally, the three sigma control limits are used, e.g., $L = 3$, but for $\lambda \leq 0.1$ Montgomery suggests using narrowed limits with $L = 2.6 \sim 2.8$ [14]. A second type of Crowder nomogram using the already determined λ value helps to find a suitable value of the parameter L [7].

2.3.4. Traditional Control Charts for Residuals of ARIMA Time Series Models

The Box–Jenkins ARIMA (autoregressive integrated moving average) time series model consists of three parts: the AR, the I, and the MA. The individual parts have specific meanings and are used in different combinations depending on the characteristics of the variable being modelled [14,64].

The AR (autoregressive) part of the ARIMA model takes into account that the variable x_t being modelled is directly dependent on previous observations $x_{t-1}, x_{t-2}, \ldots, x_{t-p}$. The autoregressive model AR (p) can be expressed using Formulas (8) or (9)

$$x_t = \xi + \phi_1 \cdot x_{t-1} + \phi_2 \cdot x_{t-2} + \cdots \phi_p \cdot x_{t-p} + \varepsilon_t, \tag{8}$$

$$\Phi(B) \cdot x_t = \xi + \varepsilon_t \tag{9}$$

where ξ and $\phi_1, \phi_2, \cdots \phi_p (-1 < \phi < 1)$ are unknown constants and B is the backshift operator. B represents a lag of one period, i.e., $B \cdot x_t = x_{t-1}$. Generally, a lag of the s period can be written using the formula $B^s \cdot x_t = x_{t-s}$. Its use simplifies the writing of models. The variable ε_t is white noise, i.e., the random component that is normally and independently distributed, the mean is equal to zero and the standard deviation is σ. Partial autocorrelation function (PACF) can be used for identification of order p for AR (p) model because is expected to "cut off" after lag p. On the other hand, the autocorrelation function (ACF) of an AR (p) process can be a mixture of exponential decay and damped sinusoid [64,84,85].

The MA (moving average) part of the ARIMA model refers to the regression error, which is a linear combination of the error terms. Their values occurred simultaneously and at different periods in the past. The moving average model MA (q) can be expressed by the formula

$$x_t = \mu + \varepsilon_t - \theta_1 \cdot \varepsilon_{t-1} - \theta_2 \cdot \varepsilon_{t-2} - \cdots - \theta_q \cdot \varepsilon_{t-q}, \tag{10}$$

or by the formula

$$x_t = \mu + (1 - \sum_{i=1}^{q} \theta_i \cdot B^i) \cdot \varepsilon_t = \mu + \Theta(B) \cdot \varepsilon_t, \tag{11}$$

where μ and $\theta_1, \theta_2, \cdots, \theta_q$ are the are unknown constants and B is the backshift operator. The correlation between x_t and previous observations $x_{t-1}, x_{t-2}, \ldots, x_{t-p}$ is equal to zero. ACF is a good tool for identifying the order q of an MA (q) process. It is expected to "cut off" after lag q. The PACF of an MA (q) process is a mixture of exponential decay and damped sinusoid [64,84,85].

The mixed autoregressive moving average model ARMA (p,q) can be expressed by combining the previous two models into the Formula (12) or expressed as Formula (13)

$$x_t = \xi + \phi_1 \cdot x_{t-1} + \phi_2 \cdot x_{t-2} + \cdots \phi_p \cdot x_{t-p} + \varepsilon_t - \theta_1 \cdot \varepsilon_{t-1} - \theta_2 \cdot \varepsilon_{t-2} - \cdots - \theta_q \cdot \varepsilon_{t-q}. \tag{12}$$

or as the formula

$$\Phi(B) \cdot x_t = \xi + \Theta(B) \cdot \varepsilon_t. \tag{13}$$

The identification of the order p and order q of the ARMA model with the help of ACF and PACF is more difficult; both exhibit exponential decay and/or damped sinusoid patterns. Therefore, Montgomery et al. suggest using an Extended sample ACF, a generalized sample PACF, or an inverse ACF [64].

Of the models described so far, AR (p), MA (q), and ARMA (p,q), are suitable for describing mainly stationary processes.

The I (integrated) part of the ARIMA model is used when the differencing is applied to the time series data x_t to reduce a non-stationary time series to a stationary one. The result is a new time series y_t, calculated according to the formula

$$y_t = (x_t - x_{t-1}) = \nabla \cdot x_t = (1 - B) \cdot x_t, \tag{14}$$

where ∇ is backward difference operator. The powers of the backshift operator B and the backward difference operator ∇ for the dth difference are defined as $\nabla^d = (1 - B)^d$. Montgomery et al. report that in most applications, first differencing (d = 1) and occasionally second differencing (d = 2) would be enough to achieve stationarity of the time series. Vandaele also points out that most series should not require more than two differences. He states that a sign of an over-differenced series is a first autocorrelation close to -0.5 and small values elsewhere [86]. The simplest non-stationary model is the ARIMA (0,1,0) model, i.e., the I(1) model, often called the random walk process [87]. It is given by the formula

$$(1 - B) \cdot x_t = \xi + \varepsilon_t. \tag{15}$$

For non-stationary processes with a variable x_t which is not fixed in the process mean, Montgomery recommends using an integrated moving average IMA (d,q) model which is given for d = 1 and q = 1 by the formula

$$(1 - B) \cdot x_t = \xi + (1 - \theta \cdot B) \cdot \varepsilon_t. \tag{16}$$

The full autoregressive integrated moving average ARIMA (p,d,q) model can be expressed by the formula

$$\Phi(B) \cdot \nabla^d \cdot x_t = \xi + \Theta(B) \cdot \varepsilon_t. \tag{17}$$

Several methods can be used to estimate the parameters of the identified ARIMA model: the method of moments, the maximum likelihood method, and the least squares method.

All calculated model parameters are first tested by the t-test to verify their significance. Then, the residuals are used for calculation, for example for an ARMA (p,q) model following the formula

$$\hat{\varepsilon}_t = x_t - \left(\hat{\xi} + \sum_{i=1}^{p} \hat{\phi}_i \cdot x_{t-i} - \sum_{i=1}^{q} \hat{\theta}_i \cdot \hat{\varepsilon}_{t-i}\right), \ t = 1, 2, \ldots, n. \tag{18}$$

Values of the root mean square error of the residuals s^2, the coefficient of determination R^2, or the adjusted R^2 can be used to compare the goodness of fit of several models. A smaller value of s^2 indicates a more appropriate model. R^2 indicates how much of variable variation is explained by the model; therefore, the higher value, the better the model. These are often expressed as percentages. The adjusted R^2 takes into account the number of model parameters, making it suitable for comparing models that differ in the number of parameters.

The suitability of the model can also be evaluated using the Akaike information criterion (*AIC*) value [88], corrected *AIC* (*AICc*) value [89], and Schwarz Bayesian information criterion (*SIC*) value [90]. These criteria penalize the sum of the square residuals for including m parameters in the model. The model that has smallest values of AIC or SIC is considered the best model. Montgomery et al. preferred using *SIC* value [64].

To assess the fit of the chosen models, the Ljung–Box test can also be used [91], which is a modification of the original Box–Pierce test. This test assesses k autocorrelations simultaneously. Therefore, it is suitable for checking whether the chosen type of ARIMA model has removed autocorrelation and thus its residuals are no longer autocorrelated.

If a suitable ARIMA model is selected and all the assumptions for its residuals are satisfied, one of the traditional control charts can be constructed.

The dynamic EWMA control chart for individual measurements is the procedure developed by Montgomery and Mastrangelo [29]. It uses the EWMA statistic with parameter smoothing parameters λ and the ARIMA (0,1,1) model that is expressed by the equation

$$x_t = x_{t-1} + \varepsilon_t - \theta_1 \cdot \varepsilon_{t-1}. \tag{19}$$

If the prediction is the value of the variable under study at time t + 1, which is calculated at the previous time step t, then

$$\hat{x}_{t+1} = EWMA_t = \lambda \cdot x_t + (1 - \lambda) \cdot EWMA_{t-1}. \tag{20}$$

The following relation (21) can calculate the one-step prediction error at time t:

$$\hat{\varepsilon}_t = x_t - \hat{x}_t, \tag{21}$$

where $\hat{x}_t$ is the estimate of the value of the variable under study at time t, which was calculated at time $t - 1$. The prediction errors follow a normal distribution and are not autocorrelated. The memory of the control chart is controlled by the size of the smoothing parameter λ, where $0 < \lambda < 1$. To determine the optimal value of the λ, an iterative procedure is recommended in which the minimum value of the sum of squares of the one-step prediction is successively found. Mastrangelo and Montgomery recommended the use of $n \geq 50$ [29]. The control limits for the dynamic EWMA control chart are calculated according to the following formulas:

$$\begin{aligned} UCL_{t+1} &= CL_{t+1} + L \cdot \hat{\sigma}_\varepsilon, \\ CL_{t+1} &= EWMA_t, \\ LCL_{t+1} &= CL_{t+1} - L \cdot \hat{\sigma}_\varepsilon\ , \end{aligned} \tag{22}$$

where $L = 3$. The estimate of the standard deviation σ_ε is calculated according to the formula

$$\hat{\sigma}_{\varepsilon} = \sqrt{\frac{\left[\sum_{t=1}^{n} \hat{\varepsilon}_t^2\right]_{opt}}{n}}, \tag{23}$$

which uses the found minimum sum of squares of the one-step prediction.

3. Results and Discussion

This section describes the calculation procedure, which includes a retrospective phase (Phase I) and a monitoring phase (Phase II) of statistical process control.

The first subsection deals with the verification of four important assumptions for the use of traditional control charts within Phase I. For this purpose, various tests and charts are used, which help to correctly identify the properties of the processed datasets of the investigated coking time variable. Based on the fulfilment or non-fulfilment of the verified assumptions, suitable control charts for monitoring the variable under study are proposed and verified in the following sections. Figure 2 shows the calculation algorithm used in this study.

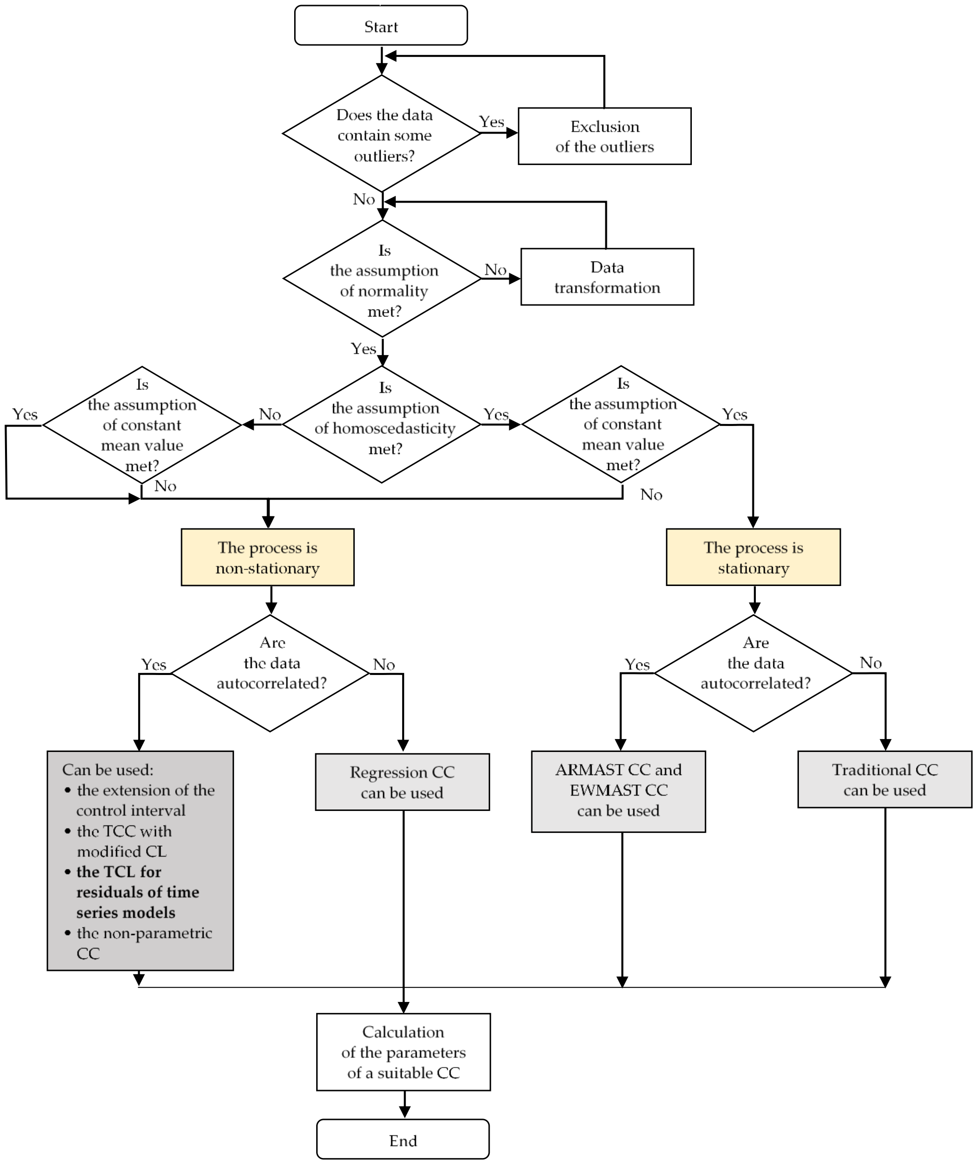


Figure 2. Procedure for selecting a suitable control chart based on assumption testing (explanatory notes: CC—control chart, TCC—traditional control chart, CL—control limit).

The next step of Phase I was to assess which of the four options identified in Section 1.4 for autocorrelated data is appropriate to use for the investigated variable. This step is described in the second subsection. All alternatives are discussed in turn and concrete arguments are given that do not allow the application of some of them.

A suitable alternative—the use of traditional control diagrams for the residuals of time series models—is elaborated in more depth. The extensive table lists all tested types of ARIMA models, together with the results of tests that confirm their suitability or inappropriateness for the monitored variable. The parameters of traditional control charts were calculated for the most suitable ARIMA (1,1,1) model at the end of Phase I. As part of this phase, the parameters of the dynamic exponentially weighted moving average control chart were also calculated. The parameters of three traditional control charts calculated for the most suitable ARIMA (1,1,1) model at the end of Phase I were subsequently used in Phase II. Similarly, in Phase II, the parameters of the dynamic control diagram of exponentially weighted moving averages, which uses the residuals of the ARIMA time series model, but whose control limits change dynamically, were also used for process monitoring.

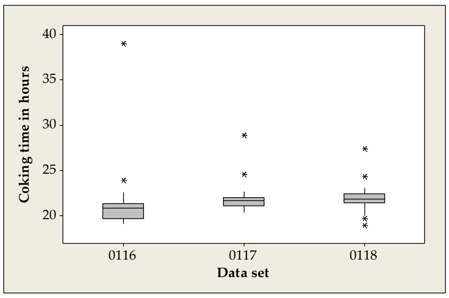
3.1. The Verification of Assumptions for the Use of Traditional Control Charts

The basic assumptions for using traditional control charts for variables are normality, homoscedasticity, constant mean value, and independent data (non-autocorrelated data). In this subsection, all these assumptions for the quality attribute of coking time will be successively verified. Table 1 contains a numerical summary of three datasets.

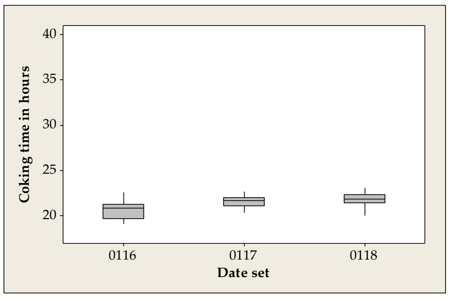
Table 1. Numerical summary of three datasets of coking time before excluding outliers.

Dataset	n	Mean	Standard Deviation	Variance	Coeff. of Variation	Skewness	Kurtosis	Min	First Quartile	Median	Third Quartile	Max	Range	Interquartile Range
0116	98	20.85	2.09	4.38	10.04	6.82	59.14	19.08	19.68	20.82	21.32	39.00	19.92	1.64
0117	98	21.68	0.99	0.98	4.58	4.14	28.31	20.28	21.08	21.65	22.02	28.87	8.58	0.94
0118	100	21.80	1.03	1.06	4.73	1.27	7.97	18.93	21.39	21.79	22.36	27.37	8.43	0.98

First, the presence of outliers in the datasets was checked. As can be seen from the box-and-whisker plots in Figure 3a, all three datasets contain outliers.



(**a**)



(**b**)

Figure 3. Box-and-whisker plots: (**a**) before excluding outliers; (**b**) after excluding outliers. ("*" indicates an outlier).

Details of the outliers are given in Table 2. In front of the lower fences on the bottom side of the box-and-whisker plot, two outliers were identified for the third dataset only. Behind the fences on the top side of the box-and-whisker plot, one outlier and one extreme outlier too were identified in every dataset. The larger number of outliers at the top of the box-and-whisker plots results from the nature of the random variable being treated—coking time. For the quality of the output product of the coke oven battery, the interval of the optimal coking time is set when the established technological prescription is observed, and the occurrence of definable causes of variability leads rather to the extension of this time above the optimal maximum.

Table 2. The identification of outliers of the coking time for three datasets.

Dataset	The Bottom Side of Box Plot				The Top Side of Box Plot			
	Far Outlier	Lower Far Fence	Outlier	Lower Fence	Upper Fence	Outlier	Upper Far Fence	Far Outlier
0116	-	14.77	-	17.23	23.78	23.88	26.24	39.00
0117	-	18.27	-	19.68	23.43	24.57	24.83	28.87
0118	-	18.46	18.93; 19.68	19.92	23.83	24.33	25.29	27.37

Eight identified outliers were excluded from the datasets. The box-and-whisker plots in Figure 3b show datasets without outliers and Table 3 contains a numerical summary of these datasets. The biggest changes in Table 3 compared to Table 1 can be seen in the reduced values of skewness, kurtosis, standard deviation, variance, and coefficient of variation.

Table 3. Numerical summary of three datasets of coking time after excluding outliers.

Dataset	n	Mean	Standard Deviation	Variance	Coeff. of Variation	Skewness	Kurtosis	Min	First Quartile	Median	Third Quartile	Max	Range	Interquartile Range
0116	95	20.64	0.92	0.84	4.45	−0.21	−0.97	19.08	19.68	20.82	21.27	22.57	3.48	1.58
0117	96	21.57	0.60	0.36	2.78	−0.04	−0.86	20.28	21.05	21.65	21.94	22.65	2.37	0.89
0118	96	21.77	0.76	0.58	3.50	−0.37	−0.76	20.03	21.40	21.79	22.35	23.03	3.02	0.95

The null hypothesis (H_0) "data follow a normal distribution" versus the alternative hypothesis (H_1) "data do not follow a normal distribution" were tested via Anderson–Darling's test, Ryan–Joiner's test, and Kolmogorov–Smirnov's test. The results of these tests are presented in Table 4. Even when outliers were excluded from the sets, the normality assumption was not always confirmed. The p-value for four executed tests is less than the chosen significance level $\alpha = 0.05$, but for five tests is more than the significance level. For dataset 0116, the null hypothesis was rejected for all three tests conducted. For the data 0117 set, the null hypothesis was not rejected for either test. For the 0118 dataset, the null hypothesis was rejected only for Anderson–Darling's test. Thus, it can be concluded that the decision on the null hypothesis being tested is not clear-cut.

Table 4. Normality test results.

Dataset	Anderson–Darling's Test			Ryan–Joiner's Test			Kolmogorov–Smirnov's Test		
	AD Statistic	p-Value	H_0	RJ Statistic	p-Value	H_0	KS Statistic	p-Value	H_0
0116	2.392	<0.005	rejected	0.969	<0.010	rejected	0.122	<0.010	rejected
0117	0.933	0.017	not rejected	0.988	0.077	not rejected	0.099	0.028	not rejected
0118	1.247	<0.005	rejected	0.982	0.015	not rejected	0.101	0.025	not rejected

If an informal approximation of the normality test, referred to as the "bold pencil test", is applied to the probability plots in Figure 4, it is possible to conclude that the coking time is a variable coming from a near-normal distribution.

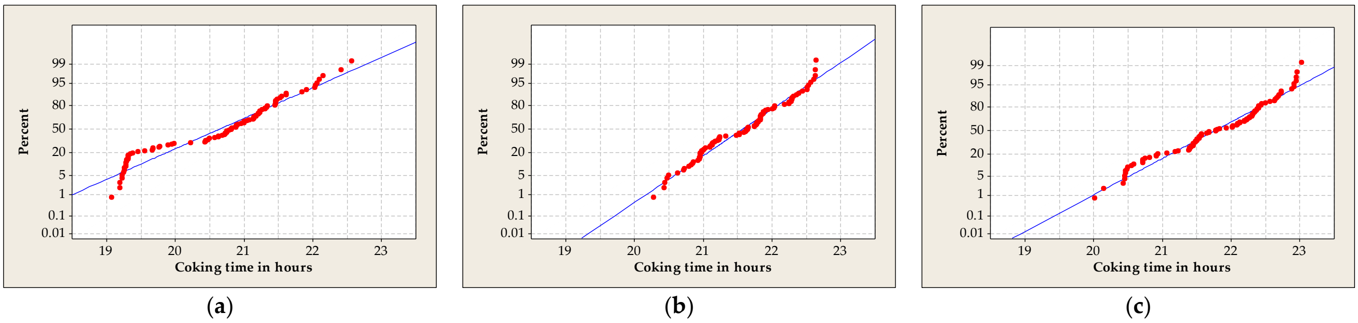


Figure 4. Probability plots for three datasets: (**a**) 0116; (**b**) 0117; (**c**) 0118.

Figure 5 shows histograms with the fitted probability density function for the three datasets. Sturges' rule was used to calculate the optimal number of classes in the histograms.

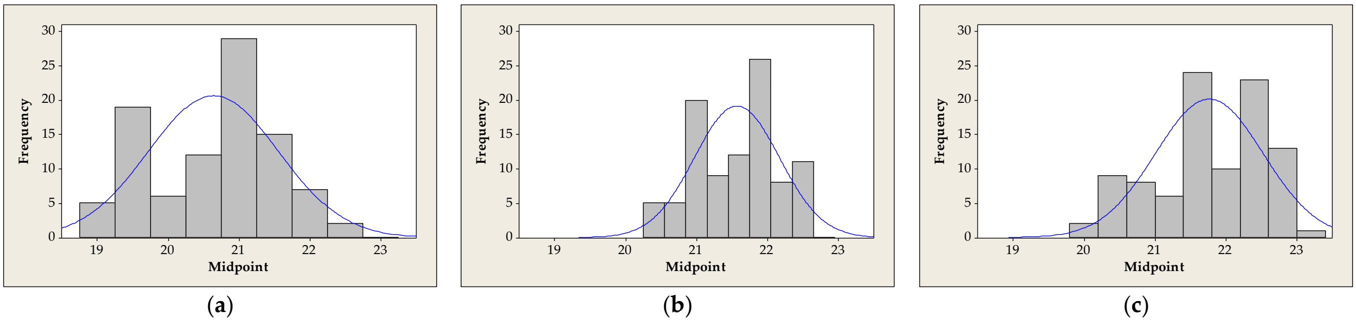


Figure 5. Histograms with fitted probability density function for three datasets: (**a**) 0116; (**b**) 0117; (**c**) 0118.

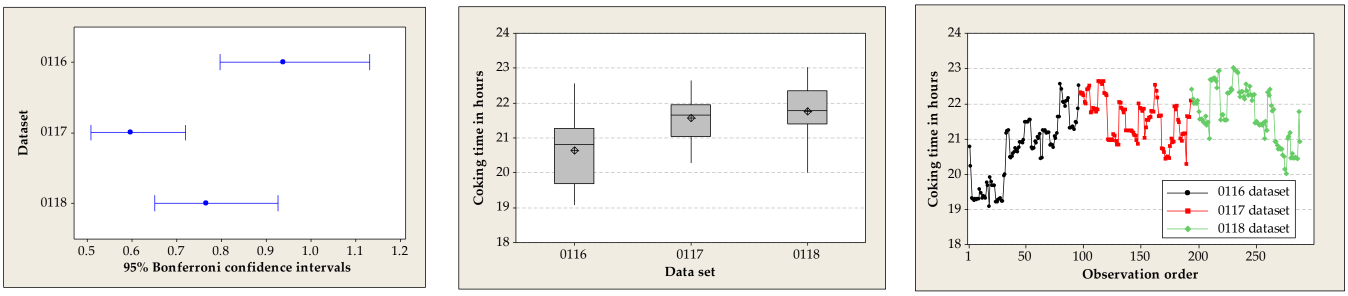
Two tests were used to test the assumption of homoscedasticity, i.e., to test the null hypothesis (H_0) "the difference between the variances of coking times is not statistically significant" versus the alternative hypothesis (H_1) "at least one variance is statistically significantly different from the others". Bartlett's test is appropriate if the data come from a normal or near-normal distribution. Levene's test is used for data from continuous but not necessarily normal distributions. The tests were applied to data from three datasets. The null hypothesis of homoscedasticity was rejected at the 0.05 significance level based on the results of both tests, as shown in Table 5. The differences in the variability of the three datasets can be seen in Figure 6a, which shows the confidence intervals for the standard deviations calculated using the Bonferroni method.

Table 5. Tests results of homoscedasticity tests and constant mean tests.

Bartlett's Test			Levene's Test			One-Way Analysis of Variance			Kruskal–Wallis' Test		
B Statistic	p-Value	H_0	W Statistic	p-Value	H_0	F Statistic	p-Value	H_0	H Statistic	p-Value	H_0
18.90	0.000	rejected	8.37	0.000	rejected	53.40	0.000	rejected	66.33	0.000	rejected

The null hypothesis (H_0) "the difference between the means of coking times is not statistically significant" versus the alternative hypothesis (H_1) "at least one mean is statistically significantly different from the others", i.e., the assumption of a constant mean of coking time over three days, was verified using a one-way analysis of variance and Kruskal–Wallis' test. This assumption was not confirmed as shown in Table 5.

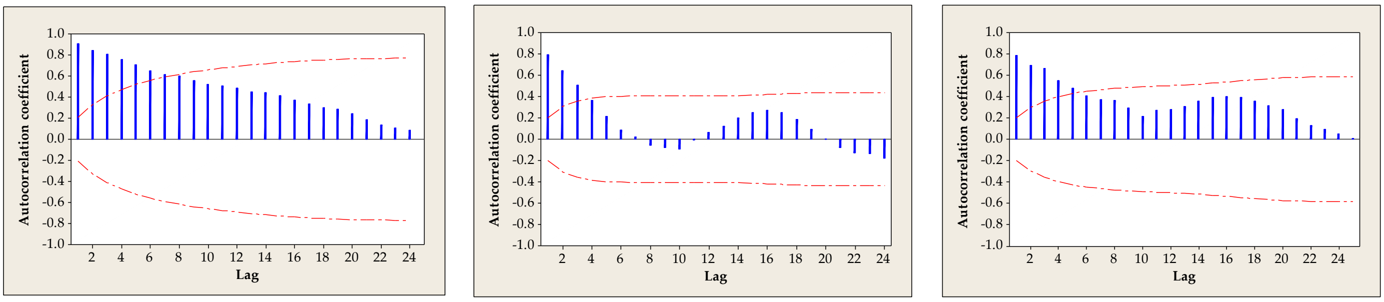
The non-constant mean coking time can be seen in the box-and-whisker plots, with the marking of mean values in Figure 6b and the time series plot in Figure 6c. It can be concluded that the coke production process is non-stationary in terms of coking time.



(**a**) (**b**) (**c**)

Figure 6. (**a**) Boniferroni confidence intervals for the three standard deviations; (**b**) the box-and-whisker plot for the three datasets; (**c**) the time series plot for the three datasets.

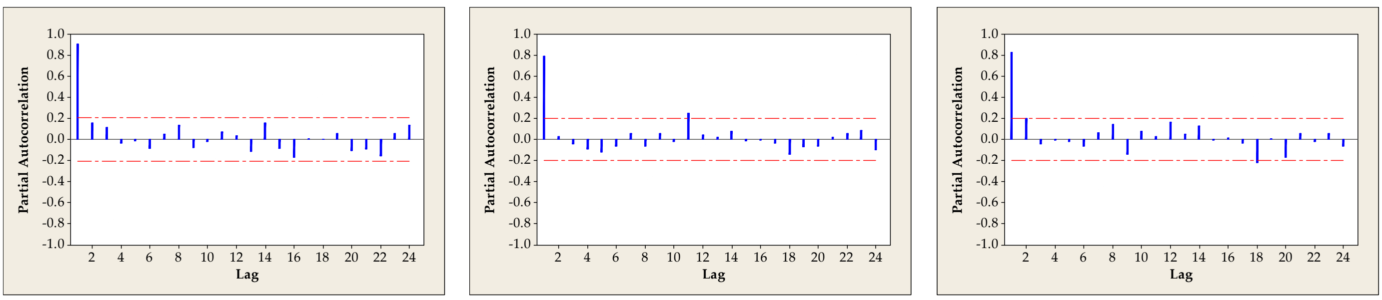
The assumption of an independent distribution of the observed variable coking time was verified using the autocorrelation coefficients. Following the general rule, autocorrelation coefficient values were calculated for 24 ($24 \leq n/4$) lags. The plots in Figure 7 show autocorrelation function plots with the 5% significance critical limits (red dashed lines) for the null hypothesis "the autocorrelations coefficients do not show statistically significant autocorrelation". Three (Figure 7b) to six (Figure 7a) autocorrelation coefficients in the plots show significant spikers because of the overrun of the critical limits. The null hypothesis must be rejected in favour of the alternative hypothesis "the autocorrelation coefficients show statistically significant autocorrelation". It can be concluded that the coking time variable exhibits a significant positive autocorrelation for all three datasets examined.



(**a**) (**b**) (**c**)

Figure 7. Autocorrelation function plots with 5% significance limits for three datasets: (**a**) 0116; (**b**) 0117; (**c**) 0118.

The slowly decreasing ACF plot in Figure 7a gives a signal that the process is non-stationary, and an opposite signal is given by the sinusoidal wave ACF in Figure 7b,c. The PACF plots (Figure 8), with a significant value at lag 1, which is close to 0.8, confirm that indeed the coking process can be deemed non-stationary.



(**a**) (**b**) (**c**)

Figure 8. Partial autocorrelation function plots with 5% significance limits for three datasets: (**a**) 0116; (**b**) 0117; (**c**) 0118.

Based on all the calculations described above, it can be concluded that coking time is a variable that does not meet three of the four assumptions for the use of traditional control charts. From Figures 6–8, it is evident that the coking time exhibits a non-stationary autocorrelated behaviour. Such a process can be monitored in several ways that have been tested and are described in the following subsections.

3.2. The Approaches to Solving Problems with Autocorrelated Coking Time Data

The following subsections describe the four identified problem-solving options using control charts for autocorrelated data, which are described in Section 1.4. It should be noted that not all the options described have a suitable application for the variable under study.

3.2.1. The Extension of the Control Interval for Monitoring of Coking Time

Due to the nature of the variable of interest, which is the coking time obtained from the individual chambers of the coke battery, it is not possible to use the former option described in Section 1.4.1. Extending the control interval would mean that not all the values obtained would be monitored, but only each k observation would be monitored, thus making the monitoring of the coke production process in the coke oven battery insufficient. This method of solving the problem with non-stationary autocorrelated processes also cannot be used because the assumptions of homoscedasticity and constant mean are not satisfied (see Section 3.1).

3.2.2. The Use of the Traditional Control Chart with Modified Control Limits

Jarošová and Noskievičová state that one of the conditions for the use of the traditional control chart with modified control limits is that the process is stationary. This condition, as mentioned above, is not fulfilled by the coking process [74]. It is a non-stationary process with a non-constant mean and heteroscedasticity.

3.2.3. The Use of the Control Charts for Residuals of Time Series Models

This option for solving problems with autocorrelated data is applicable to the variable of interest. In both phases of statistical regulation, two approaches were used, which were compared at the end.

As part of the first approach, it was necessary during Phase I to find a suitable type of ARIMA model (Section 3.2.3.1). The parameters of three traditional control charts for its residuals were then calculated and used in Phase II to monitor the coking time (Section 3.2.3.2).

The second approach—the use of a dynamic EWMA control chart—was simpler than the first one because the type of ARIMA model used results from the essence of this control chart. In this case, too, the parameters calculated in Phase I were used to monitor the coking time in Phase II.

3.2.3.1. Selection of the Most Suitable Type of ARIMA Model for the Coking Time

The procedure for selecting the most suitable type of ARIMA model is shown in Figure 9. Table 6 provides details of the ARIMA models that have been validated for the 0116 dataset data. Based on the shape of the ACF waveform in Figure 7a, which decays exponentially to zero, and the shape of the PACF waveform in Figure 8a, which has only one peak at the first position, the fit of the ARIMA (1,0,0) model was examined.

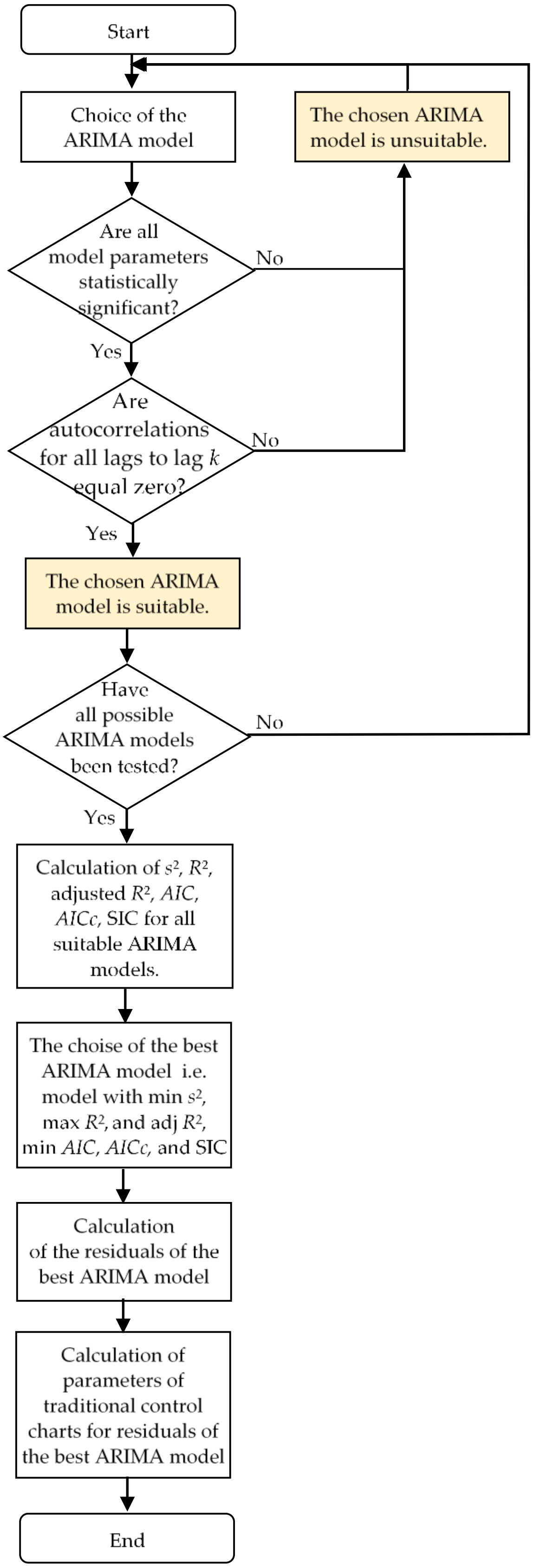


Figure 9. The procedure for selecting the most suitable type of ARIMA model whose residuals will be used in traditional control charts.

Table 6. Results of various ARIMA models for the 0116 dataset.

Type of Model	Estimated Model Parameters	Significance Test of Parameters			s^2	R^2 in %	$R^2{}_{adj}$ in %	Ljung–Box's Test for *Lag* = 24			*AIC*	AIC_C	*SIC*
		T Statistic	*p*-Value	H_0				Q Statistic	*p*-Value	H_0			
ARIMA (1,0,0)					0.142	83.47	83.59	22.6	0.425	not rejected	−181.8	−181.6	−176.5
$\hat{\xi}$	1.629	21.60	0.000	rejected									
$\hat{\phi}_1$	0.922	41.66	0.000	rejected									
ARIMA (2,0,0)					0.140	83.72	84.07	22.2	0.386	not rejected	−180.8	−181.7	−173.1
$\hat{\xi}$	1.356	34.64	0.000	rejected									
$\hat{\phi}_1$	0.787	7.57	0.160	not rejected									
$\hat{\phi}_2$	0.148	1.42	0.000	rejected									
ARIMA (0,0,1)					0.367	56.83	57.29	393.9	0.000	rejected	−93.1	−94.8	−88.0
$\hat{\mu}$	20.649	189.55	0.000	rejected									
$\hat{\theta}_1$	−0.761	−11.08	0.000	rejected									
ARIMA (1,0,1)					0.139	83.82	84.16	21.7	0.415	not rejected	−184.4	−185.2	−176.7
$\hat{\xi}$	0.991	31.61	0.000	rejected									
$\hat{\phi}_1$	0.952	25.48	0.000	rejected									
$\hat{\theta}_1$	0.199	1.79	0.077	not rejected									
ARIMA (2,0,1)					0.141	83.86	84.37	21.2	0.386	not rejected	−182.6	−184.9	−172.4
$\hat{\xi}$	0.687	29.17	0.000	rejected									
$\hat{\phi}_1$	1.168	2.19	0.031	rejected									
$\hat{\phi}_2$	−0.201	−0.40	0.689	not rejected									
$\hat{\theta}_1$	0.406	0.80	0.423	not rejected									
ARIMA (1,1,0)					0.140	83.35	83.70	23.8	0.258	not rejected	−180.6	−181.6	−172.9
$\hat{\xi}$	0.014	0.35	0.078	not rejected									
$\hat{\phi}_1$	−0.184	−1.78	0.725	not rejected									
ARIMA (0,1,1)					0.141	83.56	83.91	23.0	0.401	not rejected	−178.8	−179.9	−171.1
$\hat{\mu}$	0.012	0.42	0.674	not rejected									
$\hat{\theta}_1$	0.243	2.39	0.019	rejected									
ARIMA (1,1,1)					0.126	85.46	85.92	20.9	0.464	not rejected	−191.5	−193.4	−181.3
$\hat{\xi}$	0.007	3.62	0.000	rejected									
$\hat{\phi}_1$	0.723	8.79	0.000	rejected									
$\hat{\theta}_1$	0.973	23.34	0.000	rejected									

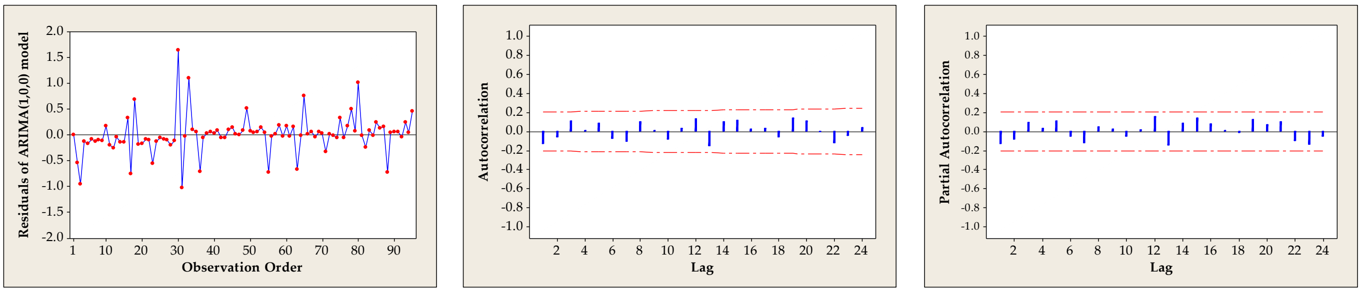
Montgomery also states that in the chemical and process industries, first-order autoregressive process behaviour is fairly common [14]. Figure 7b,c for dataset 0117 show a sinusoidal ACF waveform, and in Figure 8b,c one distinct peak at the first position and a sinusoidal waveform with a second non-synchronous peak at other positions can be seen. Therefore, the ARIMA (2,0,0) model was also tested. The ARIMA (0,0,1) model was used to demonstrate one of the inappropriate models. Next, the ARIMA (1,0,1) and ARIMA (2,0,1) models were tested. Among the recommended models for non-stationary processes, the ARIMA (1,1,0), ARIMA (0,1,1), and ARIMA (1,1,1) models were used. Model calculations were performed using Minitab 15 statistical software.

The middle part of Table 6 shows the values of the mean squared error of the residuals s^2, the coefficient of determination R^2, and the adjusted R^2 to compare the fit of the ARIMA models. The smallest value of s^2 indicates the most appropriate model. On the other hand, the highest value of adjusted R^2 indicates that the model is the best fit. The ARIMA (1,1,1) model is the best fit according to these indicators.

Table 6 also shows the results of the Ljung–Box test for k = 24. The null hypothesis of this test, "autocorrelations for all lags to lag k are zero", is tested against the alternative hypothesis, "at least one autocorrelation for lags to lag k is non-zero". The null hypothesis was rejected in favour of the alternative hypothesis only for the ARIMA (0,0,1) model.

The last three columns of Table 6 show the results of *AIC*, *AICc*, and *BIC* criteria. A model that has the smallest values of these criteria is considered the best one. According to these indicators, the ARIMA (1,1,1) model is the best fit.

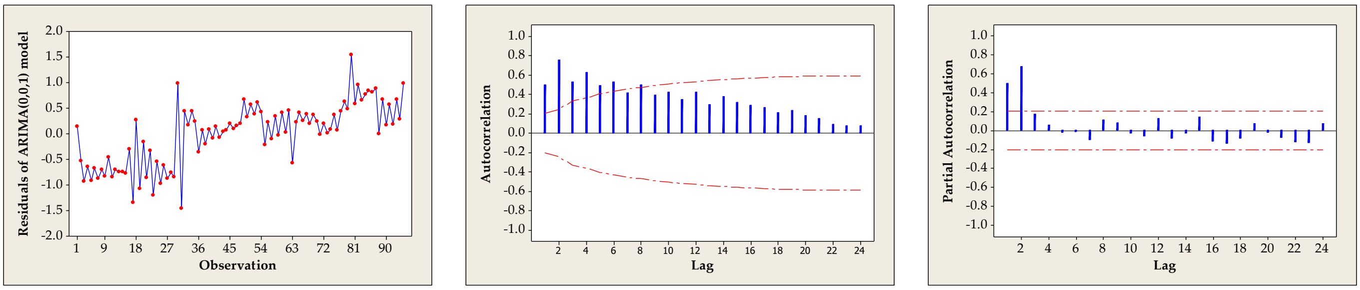
The ARIMA (1,0,0) model follows the formula $x_t = 1.629 + 0.922 \cdot x_{t-1} + \varepsilon_t$. The p-value is equal to 0.000 for both the constant $\hat{\xi}$ and the $\hat{\phi}_1$ parameter, and therefore it can be stated that both are statistically significant for $\alpha \leq 0.05$. Figure 10 shows three plots for the 0116 dataset. Figure 10a shows that the mean of the residuals is approximately stable. Figure 10b,c show a damped sinusoidal behaviour of ACF and PACF, respectively. In neither of these plots are the two standard deviation limits exceeded. These plots confirm the absence of significant autocorrelation of the residuals as well as the result of the Ljung–Box test (Q statistic = 22.6; p-value = 0.425). The $AIC = -181.8$, $AICc = -181.6$, and $SIC = -176.5$ values achieved by this model are not the lowest in the group of models studied. Based on all results presented, it can be concluded that the ARIMA (1,0,0) model could be appropriate for the variable of interest but is not the best one.



(**a**) (**b**) (**c**)

Figure 10. Plots for the ARIMA (1,0,0) model residuals for 0116 dataset: (**a**) time series plot; (**b**) autocorrelations function plot; (**c**) partial autocorrelations function plot.

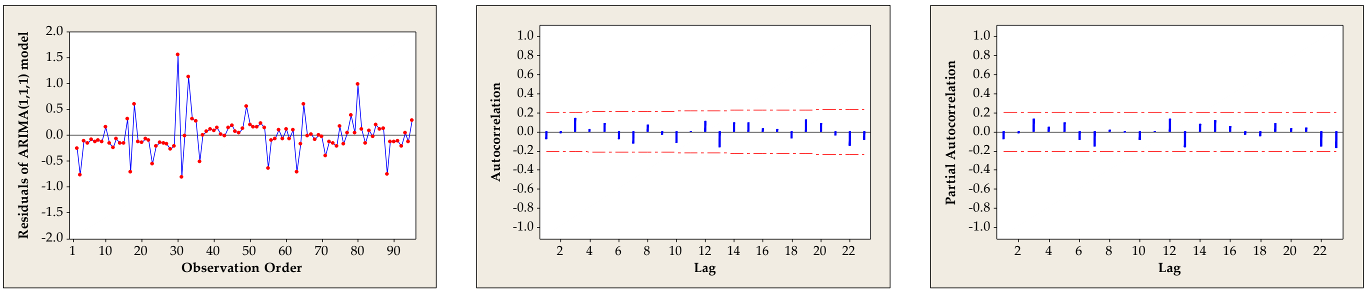
Figure 11 shows the plots for the ARIMA (0,0,1) model for the 0116 dataset. This model follows the formula $x_t = 20.649 + \varepsilon_t - 0.791 \cdot \varepsilon_{t-1}$. The p-Value is equal to 0.000 for both the constant $\hat{\mu}$ and the $\hat{\theta}_1$ parameter, and therefore it can be stated that both are statistically significant for $\alpha \leq 0.05$. In Figure 11a, it can be seen that the mean and variance of the residuals are not stable; an increasing trend is visible. The presence of significant autocorrelation is confirmed by the five values exceeding the upper red lines in the ACF plot in Figure 11b, by two values exceeding the upper red lines in the PACF plot in Figure 11c, and the results of the Ljung–Box test in Table 6 (Q statistic = 393.9; p-value = 0.000). This test suggests that autocorrelation remained in the residuals. As can be seen from Table 6, the null hypothesis was rejected in only this one case, i.e., for the ARIMA (0,0,1) model. The highest values of $AIC = -93.1$, $AICc = -94.8$, and $SIC = -88.0$ were achieved by this model, which also confirms the conclusion of its inappropriateness for the variable of interest.



(**a**) (**b**) (**c**)

Figure 11. Plots for the ARIMA (0,0,1) model for 0116 dataset: (**a**) time series plot; (**b**) autocorrelations function plot; (**c**) partial autocorrelations function plot.

The last model in Table 6 is the ARIMA (1,1,1) model that follows the formula $x_t = 0.007 + (1 + 0.723) \cdot x_{t-1} + \varepsilon_t - 0.973 \cdot \varepsilon_{t-1}$. All its parameters are statistically significant. Figure 12a shows that the mean of the residuals is stable. Figure 12b,c shows a damped sinusoidal behaviour of ACF and PACF, respectively. In neither of these plots are the two standard deviation limits exceeded. These plots confirm the absence of significant autocorrelation of the residuals as well as the result of the Ljung–Box test (Q statistic = 20.9; p-Value = 0.464). The $AIC = -191.5$, $AICc = -193.4$, and $SIC = -181.3$ values are the smallest in the group of models studied. It should be noted that ARIMA (1,1,1) is the best-fitting model of all in Table 6 and it is appropriate for the variable of interest.



(**a**) (**b**) (**c**)

Figure 12. Plots for ARIMA (1,1,1) model residuals for the 0116 dataset: (**a**) time series plot of first difference; (**b**) autocorrelations function plot; (**c**) partial autocorrelations function plot.

3.2.3.2. The Use of the Traditional Control Charts for Residuals of Selected ARIMA Models

Figure 13 shows three traditional control charts for the ARIMA (1,1,1) model residuals of dataset 0116, which was selected as the best of the compared models in Phase I. The target mean for residuals is set to 0 and the target standard deviation is set to 0.5 in all three control charts. The residuals in the control charts show a stationary pattern. Shewhart's control chart (control chart without memory) in Figure 13a for the 13th individual value signals an over-run of the upper control limit. Similarly, the CUSUM control chart (control chart with uniform and unlimited memory) signals the crossing of the upper decision interval limit (Figure 13b). The sudden shift in residuals did not cause the upper limit in the EWMA control chart (control chart with uneven and unlimited memory) to be exceeded (Figure 13c). The shift was damped by the value of the parameter λ, which was set to 0.5. Since all outliers were excluded from the 0116 dataset at the beginning of the processing prior to calculating the residuals, the EWMA control chart appears to be the best; therefore, it gives signals that the process is under control.

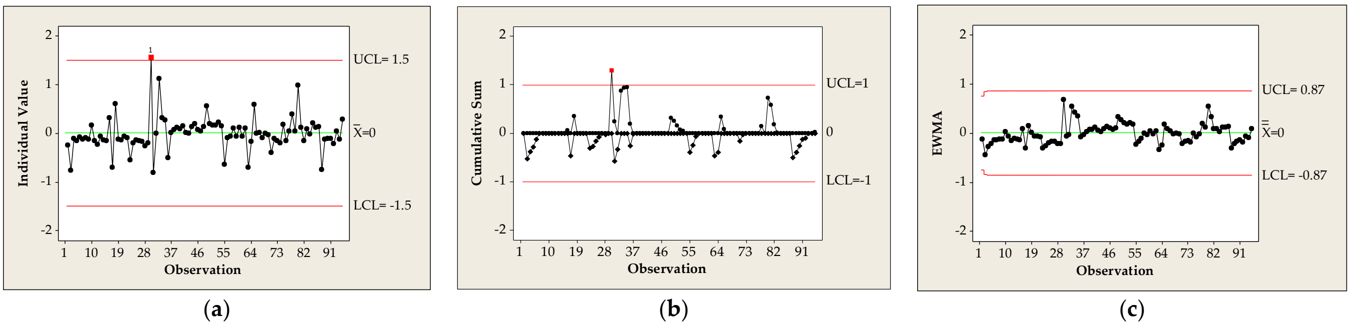


Figure 13. Traditional control charts for residuals of ARIMA (1,1,1) model of the dataset 0116: (**a**) Shewhart control chart; (**b**) CUSUM control chart; (**c**) EWMA control chart.

Subsequently, in Phase II, the parameters calculated for the 0116 dataset in Phase I. were used for the 0216 dataset obtained in February. The three control charts for the residuals of the ARIMA (1,1,1) model of the 0216 dataset are shown in Figure 14.

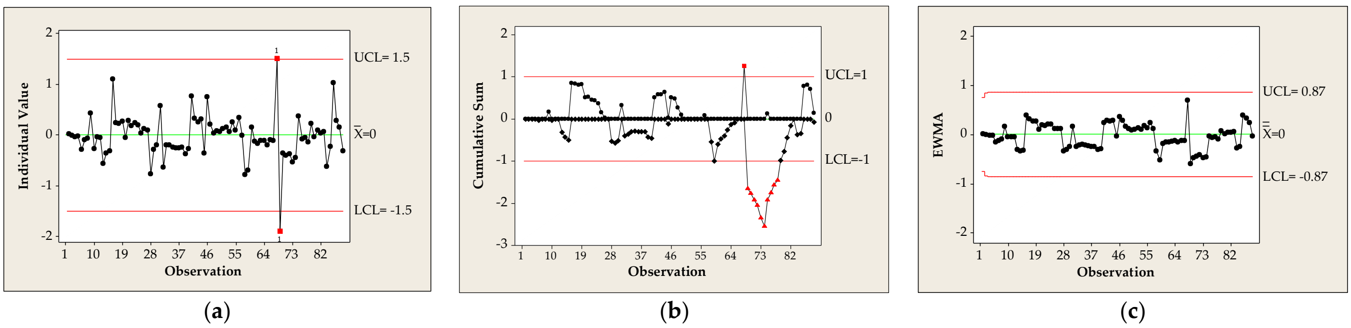


Figure 14. Traditional control charts for residuals of ARIMA (1,1,1) model of the dataset 0216: (**a**) Shewhart control chart; (**b**) CUSUM control chart; (**c**) EWMA control chart.

Shewhart's control chart (Figure 14a) detects two important signals. The 68th individual value lies just above the upper control limit and the next value lies below the lower control limit. The CUSUM control chart (Figure 14b) shows up to ten values below the lower control limit. The EWMA control chart (Figure 14c) provides no signal of exceeding the control limits.

The disadvantage of using control charts for the residuals of ARIMA models is that the scaling of the y-axis does not correspond to the measured values of the coking time. This can be inconvenient for the operator monitoring the coking process.

3.2.3.3. The Use of the Dynamic EWMA Control Chart

The dynamic EWMA control chart for individual measurements with the optimum value equal to 0.33 was calculated using the procedure described in Section 3.2.1 during Phase I. The initial $EWMA_0$ value was equal to the mean target value $\mu_0 = 22$. Figure 15 shows these control charts for the three processed datasets. The control limits change dynamically and adapt to changes in the coking time values. Since all outliers were excluded at the beginning of the processing, the control limits are not exceeded in any of the control charts, i.e., all three charts show a process that is in a state of statistical control. The calculated λ and σ_a parameters for datasets 0116, 0117, and 0118 were applied to the other three datasets obtained in February (0216, 0217, 0218) during Phase II. Outliers have not been removed from these datasets, and, as can be seen in the control charts in Figure 16, there are several cases where the dynamic control limits are exceeded.

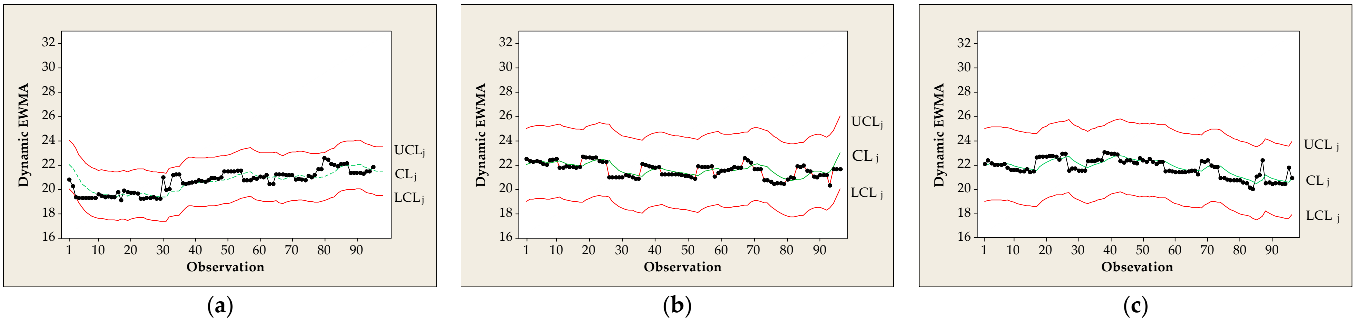


Figure 15. Dynamic EWMA control charts for individual data of the three datasets: (**a**) 0116; (**b**) 0117; (**c**) 0118.

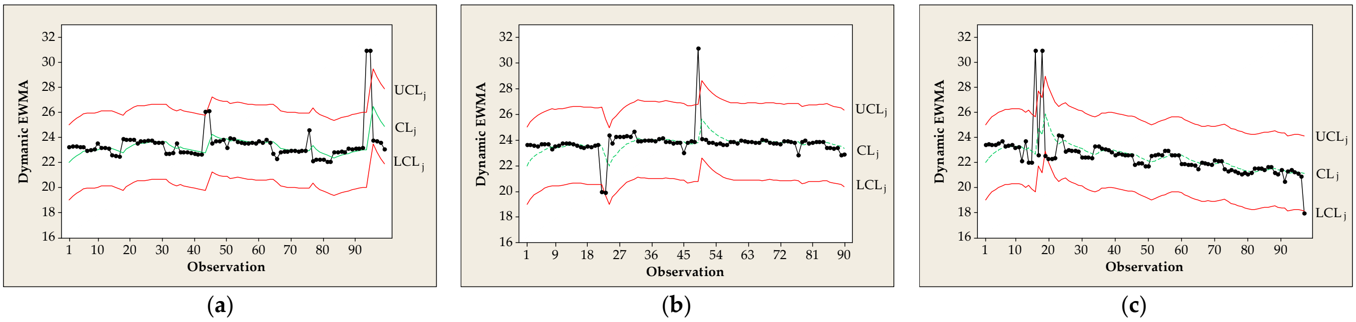


Figure 16. Dynamic EWMA control charts for individual data of the three datasets: (**a**) 0216; (**b**) 0217; (**c**) 0218.

The advantage of a dynamic EWMA control chart, as opposed to traditional control charts applied to residuals, is that it contains directly measured coking time data. The scaling of the y-axis corresponds to the unmeasured values and thus the monitoring of the coking process is more convenient for the operator than with control charts using residuals. The disadvantage of this control chart is that the control limits are not constant but change dynamically and must therefore have to be calculated for every new observation.

3.2.4. The Use of the Non-Parametric Control Chart

The main advantage of nonparametric control charts is their general flexibility since their use does not require the assumption of a particular probability distribution for the measured data to be satisfied. However, they do require the creation of samples of the measured data for every period; their application to coking time data is not possible. The individual values refer to different chambers of the coke battery and it is therefore not logical to justify their association into samples.

4. Conclusions

4.1. Summary of the Study

One of the assumptions for effective and efficient monitoring of the quality of technological processes is the use of appropriate statistical tools. Statistical process control is widely used to monitor the quality of a process or its final product. Control charts allow continuous monitoring of the process, making it easier to identify assignable causes in time so that corrective action can be taken and the process can be brought back under control.

Coking time data for three selected days in two consecutive months (January and February) were analysed. The January datasets were used to validate the assumptions about the coking time under study and to calculate the parameters of the control charts in Phase I. The February datasets were used to monitor the processes in Phase II, when control charts with calculated control limits were used.

From many existing control charts, we found a suitable type for monitoring the quality of the coke production process in the coke battery. The basic assumptions for using traditional control charts for variables are normality, homoscedasticity, constant mean, and independent data (i.e., data without autocorrelation). Multiple comparable alternatives were used in testing the assumptions to ensure that decisions were not based on a single outcome. This procedure is appropriate and recommended for similar situations. The normality of the data was verified by Anderson–Darling's test, Ryan–Joiner's test, and Kolmogorov–Smirnov's test. Through Bartlett's test, Levene's test, and Kruskal–Wallis' test, the coke production process was found to be non-stationary in terms of coking time. Based on the ACF and PACF plots, it was found that the coking time variable shows a significant positive autocorrelation. The observed data came from an approximately normal distribution but were not shown to be stable, have constant variability, and be independent. The use of traditional control charts to monitor the coking process would result in a high incidence of false signals.

For the autocorrelated variables, four possible ways forward were identified based on the literature: the extension of the control interval; the use of traditional control charts with modified control limits; the use of traditional control charts, or slightly modified versions of them, for residuals of time series models; and the use of non-parametric control charts. Based on an examination of the nature of the variable under study, the best of these options is the application of traditional control charts to the residuals of time series models.

Different types of ARIMA models have been validated and their suitability for the variable of interest has been assessed under several criteria. First, a test was used to assess the significance of the calculated parameters of the models. Next, it was determined which model had the lowest value of means square error s^2 and, conversely, the highest value of the adjusted coefficient of determination R^2. Autocorrelation was removed from the residuals due to the introduction of the autoregressive and/or integrated part of the ARIMA model, and its absence was verified using Ljung–Box's test and using the ACF and PACF plots. The calculations showed that the coking process can be evaluated as a homogeneously non-stationary autocorrelated process based on the coking time. All of the calculated criteria confirm that this model is the best fit for the variable of interest. The suitability of this model has also been demonstrated by traditional control charts applied to its residuals. The dynamic EWMA control chart proved to be even more suitable for monitoring the coking time. Its advantage is that it uses directly measured coking time data, not residuals. The scaling of the y-axis corresponds to the measured values and therefore monitoring the coking process is more convenient for the operator than with control charts using residuals.

The aim of this study was not to develop a model of the high-temperature carbonization process in the coke oven chamber, nor was it concerned with the thermal behaviour of the heating walls in the coke oven battery as in, e.g., Smolka et al. [80]. It was not the intention of the study to create a new type of control chart, nor to develop a general methodology for the application of control charts in the field. The aim was to assess and select a suitable tool for monitoring the selected control variable of the coking process from the large number of existing control chart types.

The presented EWMA dynamic control chart for individual measurements serves to monitor the coking time as one of the important control variables of the coking process. It is not intended for direct intervention in the coking process; its role is only to alert the process operator to the occurrence of problems with this control variable.

The goal of this study was not to look for complex solutions, but rather to find the simplest possible control chart that process operators could use to quickly identify the occurrence of a problem in coke production. The application of existing control charts in a new area, namely metallurgical coke production, is the main contribution of this study. Considering the lack of papers oriented to the implementation of control charts in this area, the presented approach could be an inspiration for other authors to address similar issues.

4.2. Options for Further Research Direction

Further research on monitoring the coking process could be oriented in three directions. The first one is to extend this research to other months. The second direction could make use of multivariable control charts, and the third could be based on the monitoring of profiles.

Data from 3 days of one month were used to calculate the parameters of the control charts in Phase I, and data from the next 3 days of the following month were used to verify them in Phase II. These were the winter months when coking times are longer due to the lower temperature and higher moisture content of the hard coking coal—the coke battery charge—than in the summer months. Therefore, it would be advisable to extend this research to summer months of the year when the parameters of the charge are different due to the weather.

Multivariate control charts, which have been in use since 1947, allow processes to be monitored based on the simultaneous use of multiple control variables [14]. The T^2 control chart, the multivariate EWMA control chart, and the multivariate CUSUM chart were developed by extending traditional control charts for a single variable. Their use is currently of interest because automatic control procedures allow multiple regulated variables to be measured simultaneously with relative ease. These control charts are more complex in their construction, and more difficult and complicated to interpret. A multidimensional variable used in a control chart does not give a completely unambiguous signal as to which, if any, of the original control variables are out of control. Designing an appropriate control intervention can therefore be problematic; nevertheless, it would be useful to validate this line of research in practice.

Process monitoring using control charts can be based not only on the monitoring of characteristics of quality, as used in this study, but also on the monitoring of profiles. In that case, process quality is modelled through a functional relationship between a response (dependent) variable and one or more explanatory (independent) variables. Profile monitoring, as a specialized field in SPC, emerged in the last two decades of the twentieth century. It is carried out similarly to the monitoring of characteristics of quality in two phases. Studies using different types of profiles (single and multiple linear profiles, nonlinear profiles, semiparametric and nonparametric profiles, and geometric and spatial profiles) have been published [92,93]. Certain assumptions must be verified when using profiles, as well as when using characteristics. The assumption that observations between (between-profile correlation) or within profiles (within-profile autocorrelation) are independent of each other is often violated in manufacturing practice. Several methods accounting for autocorrelation have been published [94–96]. For example, Yeganeh et al. describe the use of several machine learning techniques instead of statistical approaches in monitoring autocorrelated linear profiles [97].

The availability of suitable methods, together with technological advances, makes it possible to collect and process a large number of process control variables. This opens the way for the application of these approaches to coking process monitoring too.

Author Contributions: Conceptualization, M.B.; methodology, M.B.; software, M.B.; validation, M.B., G.B., D.B. and M.P; formal analysis, G.B., D.B. and M.B; investigation, D.B.; resources, D.B.; data curation, D.B; writing—original draft preparation, D.B. and M.B.; writing—review and editing, M.P., G.B., D.B. and M.B.; visualization, M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Slovak Research and Development Agency (APVV) under grant APVV-18-0526, by the Scientific Grant Agency (VEGA) under grant VEGA 1/0264/21, and by the Cultural and Educational Grant Agency MŠVVaŠ SR (KEGA) under grant 040TUKE-4/2021.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Montgomery, D.C. *Statistical Quality Control: A Modern Introduction*, 7th ed.; International Student Version; Wiley: Hoboken, NJ, USA, 2013; ISBN 978-1-118-32257-4.
2. Shewhart, W.A. *Economic Control of Quality of Manufactured Product*; D. Van Norstrand, Co.: New York, NY, USA, 1923.
3. Page, E.S. Control Charts for the Mean of a Normal Population. *J. R. Stat. Soc. Ser. B (Methodol.)* **1954**, *16*, 131–135. [CrossRef]
4. Lucas, J.M. The Design and Use of V-Mask Control Schemes. *J. Qual. Technol.* **1976**, *8*, 1–12. [CrossRef]
5. Hawkins, D.M. A Cusum for a Scale Parameter. *J. Qual. Technol.* **1981**, *13*, 228–231. [CrossRef]
6. Roberts, S.W. Properties of Control Chart Zone Tests. *Bell Syst. Tech. J.* **1958**, *37*, 83–114. [CrossRef]
7. Crowder, S.V. Design of Exponentially Weighted Moving Average Schemes. *J. Qual. Technol.* **1989**, *21*, 155–162. [CrossRef]
8. Crowder, S.V. A Simple Method for Studying Run-Length Distributions of Exponentially Weighted Moving Average Charts. *Technometrics* **1987**, *29*, 401. [CrossRef]
9. Lucas, J.M.; Saccucci, M.S. Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements. *Technometrics* **1990**, *32*, 1–12. [CrossRef]
10. Zhang, L.; Chen, G.; Castagliola, P. On t and EWMA t Charts for Monitoring Changes in the Process Mean. *Qual. Reliab. Eng. Int.* **2009**, *25*, 933–945. [CrossRef]
11. Shewhart, W.A.; Deming, W.E. *Statistical Methods from the Viewpoints of Quality Control. Graduate School*; The Department of Agriculture, University of California: Davis, CA, USA, 1939.
12. Qiu, P. *Introduction to Statistical Process Control*, 1st ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2013; ISBN 978-0-429-09077-6.
13. Aslam, M.; Saghir, A.; Ahmad, L. *Introduction to Statistical Process Control*, 1st ed.; Wiley: Hoboken, NJ, USA, 2020; ISBN 978-1-119-52842-5.
14. Montgomery, D.C. *Introduction to Statistical Quality Control*, 6th ed.; Wiley: Hoboken, NJ, USA, 2009; ISBN 978-0-470-16992-6.
15. Costa, A.; Fichera, S. Economic Statistical Design of ARMA Control Chart through a Modified Fitness-Based Self-Adaptive Differential Evolution. *Comput. Ind. Eng.* **2017**, *105*, 174–189. [CrossRef]
16. *ISO 7870*; Control Charts—Part 2: Shewhart Control Charts. International Organization for Standardization: Geneva, Switzerland, 2023.
17. Aziz Kalteh, A.; Babouei, S. Control Chart Patterns Recognition Using ANFIS with New Training Algorithm and Intelligent Utilization of Shape and Statistical Features. *ISA Trans.* **2020**, *102*, 12–22. [CrossRef]
18. Fuqua, D.; Razzaghi, T. A Cost-Sensitive Convolution Neural Network Learning for Control Chart Pattern Recognition. *Expert Syst. Appl.* **2020**, *150*, 113275. [CrossRef]
19. García, E.; Peñabaena-Niebles, R.; Jubiz-Diaz, M.; Perez-Tafur, A. Concurrent Control Chart Pattern Recognition: A Systematic Review. *Mathematics* **2022**, *10*, 934. [CrossRef]
20. Shu, L.; Tsui, K.; Tsung, F. Regression Control Charts. In *Wiley StatsRef: Statistics Reference Online*; Balakrishnan, N., Colton, T., Everitt, B., Piegorsch, W., Ruggeri, F., Teugels, J.L., Eds.; Wiley: Hoboken, NJ, USA, 2014; ISBN 978-1-118-44511-2.
21. Hayati, F. Regression Control Chart for Two Related Variables: A Forgotten Lesson. *IJMOM* **2017**, *6*, 262. [CrossRef]
22. *ISO 7870*; Control Charts—Part 5: Specialized Control Charts. International Organization for Standardization: Geneva, Switzerland, 2014.
23. Lu, Z.; Wang, M.; Dai, W. A Condition Monitoring Approach for Machining Process Based on Control Chart Pattern Recognition with Dynamically-Sized Observation Windows. *Comput. Ind. Eng.* **2020**, *142*, 106360. [CrossRef]
24. Cuentas, S.; García, E.; Peñabaena-Niebles, R. An SVM-GA Based Monitoring System for Pattern Recognition of Autocorrelated Processes. *Soft Comput.* **2022**, *26*, 5159–5178. [CrossRef]
25. Roberts, S.W. Control Chart Tests Based on Geometric Moving Averages. *Technometrics* **1959**, *1*, 239–250. [CrossRef]
26. Hawkins, D.M.; Olwell, D.H. *Cumulative Sum Charts and Charting for Quality Improvement*; Springer: New York, NY, USA, 1998; ISBN 978-1-4612-7245-8.
27. *ISO 7870*; Control Charts—Part 6: EWMA Control Charts. International Organization for Standardization: Geneva, Switzerland, 2016.
28. Alwan, L.C. Effects of Autocorrelation on Control Chart Performance. *Commun. Stat. Theory Methods* **1992**, *21*, 1025–1049. [CrossRef]
29. Montgomery, D.C.; Mastrangelo, C.M. Some Statistical Process Control Methods for Autocorrelated Data. *J. Qual. Technol.* **1991**, *23*, 179–193. [CrossRef]
30. Patel, A.K.; Divecha, J. Modified Exponentially Weighted Moving Average (EWMA) Control Chart for an Analytical Process Data. *J. Chem. Eng. Mater. Sci.* **2011**, *2*, 12–20. [CrossRef]
31. Cheng, S.-S.; Yang, C.-H.; Sheu, T.-S.; Yu, F.-J. On the Performance Evaluation of the Proportional Control Chart. In Proceedings of the 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Guilin, China, 29–31 July 2017; pp. 1011–1017.
32. Zhang, N.F. A Statistical Control Chart for Stationary Process Data. *Technometrics* **1998**, *40*, 24–38. [CrossRef]
33. Zhang, N.F. Statistical Control Charts for Monitoring the Mean of a Stationary Process. *J. Stat. Comput. Simul.* **2000**, *66*, 249–258. [CrossRef]
34. Woodall, W.H.; Montgomery, D.C. Research Issues and Ideas in Statistical Process Control. *J. Qual. Technol.* **1999**, *31*, 376–386. [CrossRef]

35. Shu, L.; Apley, D.W.; Tsung, F. Autocorrelated Process Monitoring Using Triggered Cuscore Charts. *Qual. Reliab. Eng. Int.* **2002**, *18*, 411–421. [CrossRef]
36. Capizzi, G.; Masarotto, G. An Adaptive Exponentially Weighted Moving Average Control Chart. *Technometrics* **2003**, *45*, 199–207. [CrossRef]
37. Testik, M.C. Model Inadequacy and Residuals Control Charts for Autocorrelated Processes. *Qual. Reliab. Eng. Int.* **2005**, *21*, 115–130. [CrossRef]
38. Castagliola, P.; Tsung, F. Autocorrelated SPC for Non-Normal Situations. *Qual. Reliab. Eng. Int.* **2005**, *21*, 131–161. [CrossRef]
39. Zou, C.; Wang, Z.; Tsung, F. Monitoring Autocorrelated Processes Using Variable Sampling Schemes at Fixed-Times. *Qual. Reliab. Eng. Int.* **2008**, *24*, 55–69. [CrossRef]
40. Khan, N.; Aslam, M.; Jun, C.-H. Design of a Control Chart Using a Modified EWMA Statistic: New EWMA Statistic. *Qual. Reliab. Eng. Int.* **2017**, *33*, 1095–1104. [CrossRef]
41. Fu, X.; Wang, R.; Dong, Z. Application of a Shewhart Control Chart to Monitor Clean Ash during Coal Preparation. *Int. J. Miner. Process.* **2017**, *158*, 45–54. [CrossRef]
42. Osei-Aning, R.; Abbasi, S.A.; Riaz, M. Optimization Design of the CUSUM and EWMA Charts for Autocorrelated Processes. *Qual. Reliab. Eng. Int.* **2017**, *33*, 1827–1841. [CrossRef]
43. Li, Y.; Li, H.; Chen, Z.; Zhu, Y. An Improved Hidden Markov Model for Monitoring the Process with Autocorrelated Observations. *Energies* **2022**, *15*, 1685. [CrossRef]
44. Yao, Y.; Chakraborti, S.; Yang, X.; Parton, J.; Lewis, D.; Hudnall, M. Phase I Control Chart for Individual Autocorrelated Data: Application to Prescription Opioid Monitoring. *J. Qual. Technol.* **2023**, *55*, 302–317. [CrossRef]
45. Reynolds, M.R.; Lu, C.-W. Control Charts for Monitoring Processes with Autocorrelated Data. *Nonlinear Anal. Theory Methods Appl.* **1997**, *30*, 4059–4067. [CrossRef]
46. Yashchin, E. Performance of CUSUM Control Schemes for Serially Correlated Observations. *Technometrics* **1993**, *35*, 37–52. [CrossRef]
47. Psarakis, S.; Papaleonida, G.E.A. SPC Procedures for Monitoring Autocorrelated Processes. *Qual. Technol. Quant. Manag.* **2007**, *4*, 501–540. [CrossRef]
48. Li, Y.; Pan, E.; Xiao, Y. On Autoregressive Model Selection for the Exponentially Weighted Moving Average Control Chart of Residuals in Monitoring the Mean of Autocorrelated Processes. *Qual. Reliab. Eng. Int.* **2020**, *36*, 2351–2369. [CrossRef]
49. Franco, B.C.; Castagliola, P.; Celano, G.; Costa, A.F.B. A New Sampling Strategy to Reduce the Effect of Autocorrelation on a Control Chart. *J. Appl. Stat.* **2014**, *41*, 1408–1421. [CrossRef]
50. Grimshaw, S.D. Constructing Control Charts for Autocorrelated Data Using an Exhaustive Systematic Samples Pooled Variance Estimator. *J. Qual. Technol.* **2023**, *55*, 349–365. [CrossRef]
51. Vasilopoulos, A.V.; Stamboulis, A.P. Modification of Control Chart Limits in the Presence of Data Correlation. *J. Qual. Technol.* **1978**, *10*, 20–30. [CrossRef]
52. Garza-Venegas, J.A.; Tercero-Gómez, V.G.; Lee Ho, L.; Castagliola, P.; Celano, G. Effect of Autocorrelation Estimators on the Performance of the X¯ Control Chart. *J. Stat. Comput. Simul.* **2018**, *88*, 2612–2630. [CrossRef]
53. Alwan, L.C.; Roberts, H.V. Time-Series Modeling for Statistical Process Control. *J. Bus. Econ. Stat.* **1988**, *6*, 87. [CrossRef]
54. Wardell, D.G.; Moskowitz, H.; Plante, P. Control Charts in the Presence of Data Correlation. *Plante* **1992**, *38*, 1084–1105. [CrossRef]
55. Wardell, D.G.; Moskowitz, H.; Plante, R.D. Run-Length Distributions of Special-Cause Control Charts for Correlated Processes. *Technometrics* **1994**, *36*, 3–17. [CrossRef]
56. Lu, C.-W.; Reynolds, M.R. EWMA Control Charts for Monitoring the Mean of Autocorrelated Processes. *J. Qual. Technol.* **1999**, *31*, 166–188. [CrossRef]
57. Jiang, W.; Tsui, K.-L.; Woodall, W.H. A New SPC Monitoring Method: The ARMA Chart. *Technometrics* **2000**, *42*, 399–410. [CrossRef]
58. Lu, C.-W.; Reynolds, M.R. Cusum Charts for Monitoring an Autocorrelated Process. *J. Qual. Technol.* **2001**, *33*, 316–334. [CrossRef]
59. Taşdemir, A. Effect of Autocorrelation on the Process Control Charts in Monitoring of a Coal Washing Plant. *Physicochem. Probl. Miner. Process.* **2012**, *48*, 495–512. [CrossRef]
60. Magaji, A.A.; Yahaya, A.; Asiribo, O.E. Assessing the Effects of Autocorrelation on the Performance of Statistical Process Control Charts. *Int. J. Math. Stat. Invent.* **2015**, *2*, 15–23.
61. Costa, A.; Fichera, S. Economic-Statistical Design of Adaptive Arma Control Chart for Autocorrelated Data. *J. Stat. Comput. Simul.* **2021**, *91*, 623–647. [CrossRef]
62. Li, Y.; Huang, M.; Pan, E. Residual Chart with Hidden Markov Model to Monitoring the Auto-Correlated Processes. *J. Shanghai Jiaotong Univ. (Sci.)* **2018**, *23*, 103–108. [CrossRef]
63. Phanthuna, P.; Areepong, Y. Detection Sensitivity of a Modified EWMA Control Chart with a Time Series Model with Fractionality and Integration. *Emerg. Sci. J.* **2022**, *6*, 1134–1152. [CrossRef]
64. Montgomery, D.C.; Jennings, C.L.; Kulahci, M. *Introduction to Time Series Analysis and Forecasting*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2016; ISBN 978-1-118-74515-1.
65. Mastrangelo, C.M.; Montgomery, D.C. SPC with Correlated Observations for the Chemical and Process Industries. *Qual. Reliab. Eng. Int.* **1995**, *11*, 79–89. [CrossRef]

66. Amin, R.W.; Reynolds, M.R.; Saad, B. Nonparametric Quality Control Charts Based on the Sign Statistic. *Commun. Stat. Theory Methods* **1995**, *24*, 1597–1623. [CrossRef]
67. Chakraborti, S.; Van Der Laan, P.; Bakir, S.T. Nonparametric Control Charts: An Overview and Some Results. *J. Qual. Technol.* **2001**, *33*, 304–315. [CrossRef]
68. Chakraborti, S.; Graham, M.A. Control Charts, Nonparametric. In *Encyclopedia of Statistics in Quality and Reliability*; Ruggeri, F., Kenett, R.S., Faltin, F.W., Eds.; Wiley: Hoboken, NJ, USA, 2007; ISBN 978-0-470-01861-3.
69. Chakraborti, S.; Eryilmaz, S. A Nonparametric Shewhart-Type Signed-Rank Control Chart Based on Runs. *Comm. Stats.—Simul. Comp.* **2007**, *36*, 335–356. [CrossRef]
70. Figueiredo, F.; Gomes, M.I. Monitoring Industrial Processes with Robust Control Charts. *REVSTAT-Stat. J.* **2009**, *7*, 151–170. [CrossRef]
71. Bakir, S.T. *A Concise Summary. Part 1, 1920′s-2000: Distribution-Free (Nonparametric) Statistical Quality Control Charts*; Createspace: Charleston, SC, USA, 2011; ISBN 978-1-4611-8743-1.
72. Koutras, M.V.; Triantafyllou, I.S. A General Class of Nonparametric Control Charts. *Qual. Reliab. Eng. Int.* **2018**, *34*, 427–435. [CrossRef]
73. Smajdorová, T.; Noskievičová, D. Analysis and Application of Selected Control Charts Suitable for Smart Manufacturing Processes. *Appl. Sci.* **2022**, *12*, 5410. [CrossRef]
74. Jarošová, E.; Noskievičová, D. *Pokročilejší Metody Statistické Regulace Procesu*, 1st ed.; Grada Publishing: Praha, Czech Republic, 2015; ISBN 978-80-247-5355-3.
75. Low, C.K.; Khoo, M.B.C.; Teoh, W.L.; Wu, Z. The Revised m -of- k Runs Rule Based on Median Run Length. *Commun. Stat. Simul. Comput.* **2012**, *41*, 1463–1477. [CrossRef]
76. Chin, W.; Khoo, M. A study of the median run length (mrl) performance of the ewma t chart for the mean. *SAJIE* **2012**, *23*, 42. [CrossRef]
77. Jones, L.A.; Champ, C.W.; Rigdon, S.E. The Performance of Exponentially Weighted Moving Average Charts with Estimated Parameters. *Technometrics* **2001**, *43*, 156–167. [CrossRef]
78. Antzoulakos, D.L.; Rakitzis, A.C. The Revised m -of- k Runs Rule. *Qual. Eng.* **2007**, *20*, 75–81. [CrossRef]
79. Díez, M.A.; Alvarez, R.; Barriocanal, C. Coal for Metallurgical Coke Production: Predictions of Coke Quality and Future Requirements for Cokemaking. *Int. J. Coal Geol.* **2002**, *50*, 389–412. [CrossRef]
80. Smolka, J.; Slupik, L.; Fic, A.; Nowak, A.J.; Kosyrczyk, L. CFD Analysis of the Thermal Behaviour of Heating Walls in a Coke Oven Battery. *Int. J. Therm. Sci.* **2016**, *104*, 186–193. [CrossRef]
81. Kucková; Fröhlichová, M. *Výroba Koksu*; Technická Univerzita v Košiciach: Košice, Slovakia, 2005; ISBN 80-8073-226-4.
82. *ISO 7870*; Control Charts—Part 4: Cumulative Sum Charts. International Organization for Standardization: Geneva, Switzerland, 2021.
83. Hawkins, D.M. Cumulative sum control charting: An underutilized spc tool. *Qual. Eng.* **1993**, *5*, 463–477. [CrossRef]
84. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice; [A Comprehensive Introduction to the Latest Forecasting Methods Using R.; Learn to Improve Your Forecast Accuracy Using Dozens of Real Data Examples]*; Otexts: Lexington, KY, USA, 2014; ISBN 978-0-9875071-0-5.
85. Box, G.E.P.; Jenkins, G.M. *Time Series Analysis: Forecasting and Control*; Holden-Day Series in Time Series Analysis and Digital Processing; Rev. ed.; Holden-Day: San Francisco, CA, USA, 1976; ISBN 978-0-8162-1104-3.
86. Vandaele, W. *Applied Time Series and Box-Jenkins Models*; Academic Press: New York, NY, USA, 1983; ISBN 978-0-12-712650-0.
87. Arlt, J.; Arltová, M. *Finanční Časové Řady*, 1st ed.; Grada: Praha, Czech Republic, 2003; ISBN 978-80-247-0330-5.
88. Akaike, H. A New Look at the Statistical Model Identification. *IEEE Trans. Automat. Contr.* **1974**, *19*, 716–723. [CrossRef]
89. Akaike, H. A Bayesian Extension of the Minimum AIC Procedure of Autoregressive Model Fitting. *Biometrika* **1979**, *66*, 237–242. [CrossRef]
90. Schwarz, G. Estimating the Dimension of a Model. *Ann. Statist.* **1978**, *6*, 461–464. [CrossRef]
91. Ljung, G.M.; Box, G.E.P. On a Measure of Lack of Fit in Time Series Models. *Biometrika* **1978**, *65*, 297–303. [CrossRef]
92. Noorossana, R.; Saghaei, A.; Amiri, A. (Eds.) *Statistical Analysis of Profile Monitoring*, 1st ed.; Wiley Series in Probability and Statistics; Wiley: Hoboken, NJ, USA, 2011; ISBN 978-0-470-90322-3.
93. Maleki, M.R.; Amiri, A.; Castagliola, P. An Overview on Recent Profile Monitoring Papers (2008–2018) Based on Conceptual Classification Scheme. *Comput. Ind. Eng.* **2018**, *126*, 705–728. [CrossRef]
94. Rahimi, S.B.; Amiri, A.; Ghashghaei, R. Simultaneous Monitoring of Mean Vector and Covariance Matrix of Multivariate Simple Linear Profiles in the Presence of within Profile Autocorrelation. *Commun. Stat. Simul. Comput.* **2021**, *50*, 1791–1808. [CrossRef]
95. Nadi, A.A.; Yeganeh, A.; Shadman, A. Monitoring Simple Linear Profiles in the Presence of Within- and Between-profile Autocorrelation. *Qual. Reliab. Eng* **2023**, *39*, 752–775. [CrossRef]

96. Khalili, S.; Noorossana, R. Online Monitoring of Autocorrelated Multivariate Linear Profiles via Multivariate Mixed Models. *Qual. Technol. Quant. Manag.* **2022**, *19*, 319–340. [CrossRef]
97. Yeganeh, A.; Johannssen, A.; Chukhrova, N.; Abbasi, S.A.; Pourpanah, F. Employing Machine Learning Techniques in Monitoring Autocorrelated Profiles. *Neural Comput. Appl.* **2023**, *35*, 16321–16340. [CrossRef]