Chapter 8

Philips case

In the previous chapters, we discussed various aspects of monitoring serially correlated data. In the present chapter, we will discuss a quality improvement project in which we ourselves were involved. Not all the topics that are brought to the attention of the reader have a direct relation with the contents of this thesis. However, in Section 8.4 it is described how monitoring serial correlation was handled in practice. We find it useful to provide a more or less complete discussion of the problems that an SPC practitioner might encounter in practice. One of these is the problem of monitoring serially correlated data with a control chart.

8.1 Introduction

Philips Semiconductors Stadskanaal is a leading supplier of diodes. Customers of Philips Semiconductors Stadskanaal are among others the automotive industry, the communications sector and manufacturers of consumer electronics. These customers are producers themselves, whose product quality is partly determined by the quality of the diodes. Therefore, Philips’ customers are demanding with respect to the quality of the diodes. They require reliable, well functioning diodes that are easy to process.

In most cases the customers solder the diodes on to a printed circuit board, so that ease of processing is to a large extent determined by solderability of the diodes. In order to ensure solderability of the diodes, Philips Stadskanaal is applying a protective tin/lead layer to the connection points of the diodes.

Insufficient layer thickness or wrong composition of the layer has been the cause of customer complaints in the past. Philips Stadskanaal is there-
fore looking for ways to improve the process of applying the tin/lead layer, the objective being a better solderability.

In the last five years, Philips Stadskanaal has acquired valuable experience in process improvement through successful application of Statistical Process Control (SPC) techniques (see Does, van Oord and Trip (1994)). The key to this success may be found in the approach that was chosen towards implementation of SPC.

At Philips Stadskanaal SPC techniques are implemented by so-called Process Action Teams (PAT’s). A PAT is constituted as follows: operators are important members because they are heavily involved in the process. The team is chaired by a responsible technical engineer, and may be complemented by a quality engineer, a service mechanic and/or a developer. A neutral outsider with profound knowledge of and experience with Statistical Process Control completes the team. A PAT receives a clear mission what to improve, and the means to realize their plans (see Does, Roes and Trip (1999)).

A PAT was started to improve the process step of tin-plating diodes. In this chapter, we will go through some of the achievements of this PAT. In the following subsections, we will introduce respectively the product, the production process, and the data that are gathered during the project. In the sections thereafter we will describe more or less chronologically the developments around this process step.

8.1.1 The product

A diode is an important electrical component possessing the special property of conducting current in only one direction, whereas it has a high resistance in the reverse direction. Diodes are used in all kinds of electrical circuits in for instance TV sets, computers, automotive ignition systems, telecommunication apparatus, power supplies for X-ray generators, and a great variety of consumer electronics.

Philips Semiconductors Stadskanaal makes four different types of glass-encapsulated diodes. Each type is made using a different manufacturing process. In this chapter we will restrict ourselves to one type of diodes, the so-called Surface Mounted Implosion Diodes (SMID’s). An exploded view of a SMID is depicted in Figure 8.1.

One of the striking things in Figure 8.1 is that the connection points of the diode are flanges rather than leads. This makes this type of diodes suitable for surface mounting, which explains the first part of the name of this type of diodes. Surface mounted diodes are smaller than leaded diodes...
and easier to process in automated industry. Philips Semiconductors Stadskanaal produces both types of diodes. However, the tin/lead layer is more critical for solderability of surface-mounted diodes than for solderability of leaded diodes, so that we will restrict ourselves to surface-mounted products.

### 8.1.2 The production process

The crystal is the heart of a diode. It is made of a small slice of the semiconducting material silicon. By contaminating both sides of a silicon wafer, one side with phosphorus, the other with boron, the silicon conducts current in one direction and blocks current in the other direction. The contaminations are brought into the silicon wafer by a diffusion process. Subsequently, several crystals are produced from one wafer.

The crystal of the diode of Figure 8.1 is placed between two studs/flange pairs. An implosion process follows that tightly fits a glass body around the crystal and the studs. The implosion process takes place in a vacuum so that the edge of the crystal only contacts the nonconductor glass. This prevents charge carriers from traveling through the diode in any other way than through the crystal. The word ‘implosion’ is part of the name of the diode in Figure 8.1 to distinguish it from diodes that are produced using other methods to apply a nonconducting body.
The studs are made of molybdenum, the flanges are made of copper. Molybdenum is a metal that has a good conductivity. When oxidized, it has the additional property that it adheres well to the glass body. The adherence of the molybdenum studs to the glass body provides mechanical strength and a hermetic sealing of the diode. The latter is very important since exposing the crystal to open air ruins the electrical properties of the diode.

Whereas oxidation of the molybdenum studs is helpful, this is not the case for oxidation of the copper flanges. One of the properties of oxidized copper is that the adherence to tin is bad. The customers of the type of diodes of Figure 8.1 solder the diodes with tin, so that solderability is adversely influenced by the oxidation of the copper flanges. To overcome this problem, the oxidized copper is removed, and a tin/lead layer that prevents the flanges from coming into contact with oxygen is applied. In the remainder of this chapter, we will focus on this production step.

After tin-plating, a 100% inspection follows. Several electrical characteristics of each diode are measured, and it is decided to accept or to reject the diode. For some types of diodes, a classification in sub-types is made. The tin/lead layer is not only important for solderability but also facilitates the inspection. If it were absent, a potential difference would arise between copper-oxide on the surface of the flange and the pure copper inside the flange. This would introduce a large measurement error.

In the following process step, a white band is painted on the glass body of the accepted diodes (see Figure 8.1). The placement of the band indicates the cathode side of the diode. The electrical characteristics can be deduced from a code that is also printed on the glass body.

As a final process step, the diodes are packed in such a way that the customers can process the diodes in an automated way.

8.1.3 Relevant quality characteristics of the tin/lead layer

In the remainder of this chapter, we confine ourselves to only one of the process steps, namely applying a tin/lead layer. As discussed in the previous subsection, the tin/lead layer is important for the customers, since it determines the solderability of the diodes. The layer also facilitates inspection, which is the process step directly following the tin-plating of the diodes. The quality of the tin/lead layer is determined by two characteristics: solderability and composition.

The first quality characteristic is difficult to ascertain, since it is the net result of a complex of determinants such as thickness of the layer,
composition of the layer, pollution of the layer with organic materials, and the like. Therefore, “a good solderability” is translated into a requirement on a more ascertainable characteristic of the layer: viz. layer thickness.

If the diodes were soldered right after production, any thickness of the tin/lead layer would provide sufficient protection against oxidation. However, the diodes are shipped to factories and warehouses all over the world, where they may be kept in stock for some time. If the thickness of the tin/lead layer is not sufficient, diffusion of copper atoms in the tin/lead layer will result in copper atoms reaching the surface of the layer, and these atoms will oxidize. This phenomenon is one of the causes of bad solderability. For this reason, a lower specification limit is set for the thickness of the layer. A minimal thickness of 1\(\mu\)m \((1 \times 10^{-6}\text{m})\) proves to be sufficient to warrant good solderability after a two years stay in any warehouse (provided certain conditions concerning relative humidity and temperature are met).

Excessive thickness may cause problems during the 100% inspection in the one but last step of the production. Diodes with a tin/lead layer that is too thick (say more than 50\(\mu\)m) will not pass the sieve that is used to prevent crooked diodes from entering the measuring apparatus. However, a strict upper specification limit is not used in practice, and will therefore not be considered. The target thickness of the tin/lead layer is 10\(\mu\)m.

The second quality characteristic, the composition of the layer, concerns the ratio of tin and lead in the tin/lead layer. This characteristic is directly observable on the products. If the portion of lead in the tin/lead layer is too large, the layer will be too soft. Putting a diode with too soft a layer through the measuring apparatus will cause some of the tin/lead layer to be worn off. These grindings smudge the glass body of subsequent diodes, which may lead to leakage of current over the body. Such diodes are rejected, although the electrical characteristics may have been perfect before entering the measuring apparatus. To guard against waste caused by this phenomenon, the lower specification limit of the tin portion in the tin/lead layer is set to 77%. The target value is 80%. In the next section we will describe how the tin/lead layer is applied.

### 8.1.4 The tin-plating process

The protective layer is applied to the diode by electrogalvanizing the connection points with a tin/lead alloy. A schematic view of the tin-plating process is depicted in Figure 8.2.

The tin-plating takes place in a conducting chemical bath. A rotating
Figure 8.2: A schematic view of the tin-plating process.

basket that is permeable with respect to the tin-plating liquid is filled with diodes and placed into the bath. The anode side of the voltage supply consists of a 80% tin, 20% lead alloy that is gradually dissolving in the bath liquid. The dissolved metal ions precipitate on diodes contacting the cathode side of the voltage supply. That is, on the cathode side we have the following chemical reactions:

\[
\text{Sn}^{2+} + 2e^- \rightarrow \text{Sn}
\]

and

\[
\text{Pb}^{2+} + 2e^- \rightarrow \text{Pb},
\]

where Sn is the chemical symbol for tin, Pb is the chemical symbol for lead, and electrons are denoted by \(e^-\).

As a result of this reaction on the cathode side of the voltage supply, the concentration of metal ions in the liquid will decrease. By using an anode of a 80% tin, 20% lead alloy (see Figure 8.2) it is possible to replenish the metal ions. With current running through the bath, electrons of tin and lead molecules in the anode are sent away into the direction of the voltage supply and the remaining metal ions dissolve in the liquid. Hence, the reverse of the chemical reaction that takes place at the cathode will
take place at the anode. The net effect is that tin and lead molecules are transferred from the anode of the voltage supply to diode flanges.

Important parameters of the process that influence the quality of the tin/lead layer are among others the temperature of the bath, the current density, the length of the time span the diodes stay in the bath, the voltage and the chemical composition of the bath. In order to maintain a high quality of the tin/lead layer, controlling each of these parameters is required. In practice it turned out to be very difficult to devise a proper control strategy that could deal with exhaustion of the bath due to the production process. This is the problem the PAT focused on. For the combined effect of other parameters we will assume that they do not have a systematic effect on the tin/lead layer.

The chemical composition of the tin-plating bath changes over time. This is the result of exhaustion of the bath due to running the production process, and evaporation of the tin-plating liquid. In order to maintain a good quality of the tin/lead layer, the concentrations of the components of the bath must fall between certain limits. The bath is a mixture of the following five components:

1. acid;
2. an \( \text{Sn}^{2+} \) solution;
3. a \( \text{Pb}^{2+} \) solution;
4. brightener;
5. formalin.

The acid is thought of as an important component since it takes care of the conductivity of the bath. The two metal solutions are also considered to be important since the \( \text{Sn}^{2+} \) and \( \text{Pb}^{2+} \) ions present in them precipitate on the diodes to form the tin/lead layer. The remaining components, brightener and formalin, were thought to play a certain part in the tin-plating process. However, it turned out that their importance was underrated (see Subsection 8.3.4). The brightener takes care of a smooth layer, whereas the formalin helps to control the proportions of tin and lead in the resulting layer.

During the production process, the composition and the volume of the tin-plating liquid changes. Changes in the composition are due to chemical reactions that take place in the production process. Furthermore, some components (e.g. formalin) evaporate faster than others, resulting in a
change in their relative proportions. The overall volume of the bath decreases due to evaporation and to the dragging out of the tin-plating liquid together with the diodes. As a result, the tin-plating bath needs to be replenished regularly.

For each of the components, a lower limit is set on the concentration. If the concentration of one of the components falls below its limit, the quality of the tin/lead layer deteriorates. Likewise, on the concentrations of acid, $\text{Sn}^{2+}$ solution, and $\text{Pb}^{2+}$ solution an upper limit is set.

If the composition of the bath is such that one or more of the concentrations fall outside the limits, action is required. Fortunately, highly concentrated solutions for each of the components are available so that concentrations can be raised by adding an appropriate mix of these solutions to the bath. The bath can be diluted using demineralized water.

Before the PAT was started the bath used to be replenished with fixed additions of the separate components, independent of the actual composition of the bath. This replenishment strategy was not satisfactory, since it resulted in a high level of chemicals being used, whereas the bath did not appear to be stable. The instability of the bath affected the quality of the tin/lead layers on the diodes in such a way that customer complaints were received. A new replenishment strategy was, therefore, adopted that used the actual contents of the bath to determine the additions. To this end, the bath was analyzed every day. Based on the results of the analysis, a computer program determined what additions were necessary to ensure that the bath was fit for production again. The new strategy resulted in a large reduction in the amounts of chemicals being used. The quality of the tin/lead layer however remained unsatisfactory.

Therefore, further improvement was necessary. At this point the author got involved in the problem. The first improvement he was able to suggest concerned the way in which the additions were computed. In the next section this problem is formulated as a Linear Programming (LP) problem, for which an optimal solution (in the sense of minimal costs of additions) can be found.

## 8.2 Computing additions to the tin-plating bath

### 8.2.1 Introduction

In the previous section it was discussed that the chemical composition of the tin-plating bath changes over time. In order to maintain a good quality of the tin/lead layer, the concentrations of the components of the bath
must fall between certain limits. If the composition of the bath is such that one or more of the concentrations fall outside the limits, action is required. Fortunately, there are highly concentrated solutions for each of the components which can be added in case its concentration has become too low. If the concentration of one of the components is too high, demineralized water can be added to lower the concentration.

So far the problem seems quite straightforward: concentrations that are too low can be raised by adding some of the appropriate solutions. Concentrations that are too high can be lowered by thinning the bath. However, computing how much is needed of every solution is not as straightforward as it may seem at first sight. The problem is that all five concentrations change if one of the components is added. This may cause trouble if we compute the additions for each of the components one by one. This can be illustrated by means of an example.

Suppose that we have computed how much of the first four components must be added so that the requirements are met. Then it may happen that the addition of the fifth component lowers the concentrations of the first four in such a way that some of these fall below their lower limits. This means that an additional amount of such components is required. But these additions in their turn lower the concentrations of all other components, which again may lead to extra additions, etc.

Obviously, this “one-by-one” strategy brings with it the danger of recommending large additions, even in situations where it would have been possible to recondition the bath using smaller additions. Furthermore, the strategy completely ignores the costs of the additions to be made. In this section, the replenishment problem is formulated as a Linear Programming model that has minimization of the costs of the additions as its objective, whereas requirements with regard to the concentrations serve as restrictions. The input of the model is a chemical analysis of the bath, and the output is a prescription of how much of each component must be added so that the bath meets all requirements. The output will be an optimal solution to the replenishment problem in the sense that it is not possible to find a cheaper set of additions that will bring all the concentrations within their limits.

Minimization of costs of the additions is not the only advantage of this approach. As a result of too large additions, it regularly happened that tin-plating liquid had to be drained off to make place for the recommended additions. However, the liquid contains heavy metals, and must be purified before it is allowed to be drained off. Hence, besides the high cost associated with adding too much of expensive chemicals to the bath, additional costs
are charged by the purifying department for draining off the surplus of tin-plating liquid. The Linear Programming approach discussed in this section allows us to take these costs into account, too. The practical examples that are discussed at the end of this section show that draining off old tin-plating liquid can be avoided in some cases. Besides cost reduction, this is also beneficial for the environment.

This section is organized as follows: in Subsection 8.2.2, we will introduce the mathematical notation that is used in the remainder of this section; this in its turn will be followed by a mathematical formulation of the problem in Subsection 8.2.3. The Simplex algorithm will be used to solve the Linear Programming model for which a starting point is required. In Subsection 8.2.4, a standard starting point is derived that can be used for every instance of the problem. In Subsection 8.2.5 we will compare some of the results of our model to the results of a computer program that was used previously at Philips Semiconductors Stadskanaal. A short conclusion in Subsection 8.2.6 ends this section.

### 8.2.2 Preliminaries

In Table 8.1, an overview is presented of the mathematical notation that will be used in the remainder of this section.

<table>
<thead>
<tr>
<th>component</th>
<th>observed concentration</th>
<th>lower, upper limit</th>
<th>addition in liters</th>
<th>concentration of addition</th>
<th>cost of addition</th>
</tr>
</thead>
<tbody>
<tr>
<td>acid</td>
<td>$m_1$ (gr/l)</td>
<td>$l_1, u_1$</td>
<td>$x_1$</td>
<td>$d_1$ (gr/l)</td>
<td>$c_1$ (Dfl/l)</td>
</tr>
<tr>
<td>Sn$^{2+}$ solution</td>
<td>$m_2$ (gr/l)</td>
<td>$l_2, u_2$</td>
<td>$x_2$</td>
<td>$d_2$ (gr/l)</td>
<td>$c_2$ (Dfl/l)</td>
</tr>
<tr>
<td>Pb$^{2+}$ solution</td>
<td>$m_3$ (gr/l)</td>
<td>$l_3, u_3$</td>
<td>$x_3$</td>
<td>$d_3$ (gr/l)</td>
<td>$c_3$ (Dfl/l)</td>
</tr>
<tr>
<td>brightener</td>
<td>$m_4$ (gr/l)</td>
<td>$l_4$</td>
<td>$x_4$</td>
<td>$d_4$ (ml/l)</td>
<td>$c_4$ (Dfl/l)</td>
</tr>
<tr>
<td>formalin</td>
<td>$m_5$ (gr/l)</td>
<td>$l_5$</td>
<td>$x_5$</td>
<td>$d_5$ (gr/l)</td>
<td>$c_5$ (Dfl/l)</td>
</tr>
<tr>
<td>demin. water</td>
<td></td>
<td></td>
<td>$x_6$</td>
<td></td>
<td>$c_6$ (Dfl/l)</td>
</tr>
<tr>
<td>drain off</td>
<td></td>
<td></td>
<td>$x_7$</td>
<td></td>
<td>$c_7$ (Dfl/l)</td>
</tr>
<tr>
<td>acid in Sn$^{2+}$ sol.</td>
<td></td>
<td></td>
<td></td>
<td>$d_*$ (gr/l)</td>
<td></td>
</tr>
</tbody>
</table>

As the input of our model, we have the measurements of the concentrations of the bath components. The observed concentration of component $i$, ($i = 1, \cdots, 5$), is denoted by $m_i$. 
8.2. COMPUTING ADDITIONS TO THE TIN-PLATING BATH

For all of the components, a lower limit is set on the concentration. On the concentration of acid, \( \text{Sn}^{2+} \) solution, and \( \text{Pb}^{2+} \) solution also an upper limit is set. We denote the lower limit for component \( i \) by \( l_i \) and the upper limit for component \( i \) by \( u_i \).

Highly concentrated solutions are available for each of the five components. These solutions can be added in order to raise the corresponding concentrations in the tin-plating liquid. This results in five decisions: for each component, we have to determine the size of the addition, thereby taking into account that, after replenishment, all requirements on the concentrations must be met. With decision \( i \) we associate a decision variable \( x_i \) that indicates the number of liters we add from component \( i \) for \( i = 1, \ldots, 5 \).

If the concentration of one of the components is too high, demineralized water can be added. This calls for a sixth decision variable \( x_6 \) indicating the number of liters of demineralized water we decide to add. Furthermore, in cases where the bath volume is high, some tin-plating liquid may have to be drained off before making the additions associated with \( x_1, x_2, \ldots, x_6 \) in order to prevent the bath from overflowing. To this end, we define \( x_7 \) as the number of liters we decide to drain off before making the additions. We emphasize ‘before’ because the number of liters to be drained off was not taken into consideration in the old situation. The new bath volume was reported, whether the bath was overflowing or not. So, after the additions were computed it was known how many liters had to be drained off. This is in general not equal to the number of liters to be drained off before making additions, since draining off affects the additions to be made.

The concentrations of the highly concentrated solutions are also needed in the computations. We denote them by \( d_1, \ldots, d_5 \). Acid is available as a separate component. However, it is also the solvent of the \( \text{Sn}^{2+} \) solution. The mathematical model that will be developed in the next section has to take into account that with every addition of \( \text{Sn}^{2+} \) solution, acid will be added to the bath, too. We let \( d_s \) denote the concentration of acid in the \( \text{Sn}^{2+} \) solution.

With each of the decision variables, costs are associated. For the first six variables this is simply the price per liter, whereas for the cost associated with \( x_7 \) the cost price of purifying one liter of tin-plating liquid is used. We denote these costs coefficients by \( c_1, \ldots, c_7 \).

8.2.3 Mathematical formulation of the problem

Our objective is to minimize the costs of replenishing the bath, under the restriction that the concentrations fall between predetermined limits. With
the symbols introduced in the preceding section we can express our objective as follows

\[ \text{minimize } c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 + c_6x_6 + c_7x_7. \]

It is also possible to formulate the restrictions on the concentrations in mathematical terms. If we let \( V_{\text{old}} \) denote the volume of the bath before replenishing, our new bath volume will be \( V_{\text{old}} + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 - x_7 \). Since the new bath volume may not exceed a certain maximum bath volume \( b_{\text{max}} = 600 \) liters, we have as first constraint

\[ V_{\text{old}} + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 - x_7 \leq b_{\text{max}}. \]

The new bath volume must exceed a certain minimal value \( b_{\text{min}} = 450 \) liters, leading to the following constraint:

\[ V_{\text{old}} + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 - x_7 \geq b_{\text{min}}. \]

The tin-plating liquid contains \( m_2(V_{\text{old}} - x_7) \) grammes of tin directly after an amount of \( x_7 \) liters is drained off and just before any additions are made. We make the decision to add \( d_2x_2 \) grammes of tin, so that after replenishment the bath contains \( m_2(V_{\text{old}} - x_7) + d_2x_2 \) grammes of tin. If we divide this by the new bath volume, we have the concentration of tin after replenishment. Since this concentration must exceed \( l_2 \), we arrive at the following constraint

\[ \frac{m_2(V_{\text{old}} - x_7) + d_2x_2}{V_{\text{old}} + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 - x_7} \geq l_2 \]

which can be rewritten as

\[ l_2x_1 + (l_2 - d_2)x_2 + l_2x_3 + l_2x_4 + l_2x_5 + \]
\[ + l_2x_6 + (m_2 - l_2)x_7 \leq (m_2 - l_2)V_{\text{old}}. \]

Analogously, we find the following constraint associated with the upper limit of the tin concentration in the bath

\[ -u_2x_1 + (d_2 - u_2)x_2 - u_2x_3 - u_2x_4 - u_2x_5 - u_2x_6 + \]
\[ + (u_2 - m_2)x_7 \leq (u_2 - m_2)V_{\text{old}}. \]
Following exactly the same line of reasoning, we find constraints associated with the lower and upper limit of the concentration of lead, and a constraint associated with the lower limit of the concentration of brightener.

So far, we skipped the constraints associated with limits of the concentration of acid and formalin. The reason for this is that these constraints are different from the constraints above.

For the constraints on the concentration of acid, the difference is due to the fact that there are two ways to add acid to the bath. First, like all other components, we have acid available as a separate component for making additions to the tin-plating liquid. But secondly, acid is also the solvent of the Sn$^{2+}$ solution, so that with each addition of Sn$^{2+}$ solution, acid is added to the bath, too. Therefore, besides the addition of acid, which we denoted by $x_1$, we have to take the amount of Sn$^{2+}$ solution added ($x_2$) into account when controlling the concentration of acid.

The number of grammes of acid we decide to add to the tin-plating liquid can be written as $m_1(V_{old} - x_7) + d_1x_1 + d_2x_2$. Just before any additions are made, $m_1(V_{old} - x_7)$ grammes of acid are present in the bath, so that after addition we can write the concentration of the acid as

$$\frac{m_1(V_{old} - x_7) + d_1x_1 + d_2x_2}{V_{old} + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 - x_7}.$$

The lower limit on the concentration of acid then leads to the following constraint

$$(l_1 - d_1)x_1 + (l_1 - d_2)x_2 + l_1x_3 + l_1x_4 + l_1x_5 +$$

$$+ l_1x_6 + (m_1 - l_1)x_7 \leq (m_1 - l_1)V_{old},$$

whereas the upper limit can be expressed as

$$(d_1 - u_1)x_1 + (d_1 - u_1)x_2 - u_1x_3 - u_1x_4 - u_1x_5 - u_1x_6 +$$

$$+ (u_1 - m_1)x_7 \leq (u_1 - m_1)V_{old}.$$

The constraint associated with the lower limit of formalin differs from other constraints due to the fact that no measurement device is available to determine $m_5$, the concentration of formalin. In practice, the concentration of formalin is estimated at its lower limit and enough formalin is added to ensure that its concentration does not fall below the lower limit if additions of the other components are made.
If the concentration of formalin were observable we would have had the constraint
\[ l_5x_1 + l_5x_2 + l_5x_3 + l_5x_4 + (l_5 - d_5)x_5 \\
+ l_5x_6 + (m_5 - l_5)x_7 \leq (m_5 - l_5)V_{old}, \]
but by setting \( m_5 = l_5 \), this reduces to
\[ l_5x_1 + l_5x_2 + l_5x_3 + l_5x_4 + (l_5 - d_5)x_5 + l_5x_6 \leq 0. \]

Note that this constraint is independent of \( V_{old} - x_7 \), the volume after draining off and before adding. It only depends on the volume of the additions. If the unknown concentration of formalin satisfies \( m_5 > l_5 \), this constraint is more restrictive than the constraint we would have had if \( m_5 \) was available, and less restrictive if \( m_5 < l_5 \). For this reason, combined with the fact that formalin evaporates quickly, the process engineers prescribed a minimal addition of 0.5 liters of formalin. Hence, we include the following constraint in our model
\[ x_5 > 0.5. \]

To complete the model, a few more constraints must be added. The first is that it is not possible to drain off more tin-plating liquid than the old volume of the bath:
\[ x_7 \leq V_{old}. \]

And finally, we cannot add or drain off negative amounts. That is, we add the following nonnegativity constraints
\[ x_i \geq 0 \quad \text{for} \ i = 1, 2, \ldots, 7. \]

We have formulated the replenishment problem as a problem of minimization of a linear objective function, subject to linear constraints. Summarized, the complete model is:
8.2. COMPUTING ADDITIONS TO THE TIN-PLATING BATH

\[
\begin{align*}
\text{min} & \quad c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_4 + c_5 x_5 + c_6 x_6 + c_7 x_7 \\
\text{s.t.} & \quad -x_1 - x_2 - x_3 - x_4 - x_5 - x_6 + x_7 \leq V_{\text{old}} - b_{\text{min}} \\
& \quad x_1 + x_2 + x_3 + x_4 + x_5 + x_6 - x_7 \leq b_{\text{max}} - V_{\text{old}} \\
& \quad (l_1 - d_1)x_1 + (l_1 - d_1)x_2 + l_1 x_3 + l_1 x_4 + l_1 x_5 + l_1 x_6 + (m_1 - l_1)x_7 \leq (m_1 - l_1)V_{\text{old}} \\
& \quad (d_1 - u_1)x_1 + (d_1 - u_1)x_2 - u_1 x_3 - u_1 x_4 - u_1 x_5 - u_1 x_6 + (u_1 - m_1)x_7 \leq (u_1 - m_1)V_{\text{old}} \\
& \quad l_2 x_1 + (l_2 - d_2)x_2 + l_2 x_3 + l_2 x_4 + l_2 x_5 + l_2 x_6 + (m_2 - l_2)x_7 \leq (m_2 - l_2)V_{\text{old}} \\
& \quad -u_2 x_1 + (d_2 - u_2)x_2 - u_2 x_3 - u_2 x_4 - u_2 x_5 - u_2 x_6 + (u_2 - m_2)x_7 \leq (u_2 - m_2)V_{\text{old}} \\
& \quad l_3 x_1 + l_3 x_2 + (l_3 - d_3)x_3 + l_3 x_4 + l_3 x_5 + l_3 x_6 + (m_3 - l_3)x_7 \leq (m_3 - l_3)V_{\text{old}} \\
& \quad -u_3 x_1 - u_3 x_2 + (d_3 - u_3)x_3 - u_3 x_4 - u_3 x_5 - u_3 x_6 + (u_3 - m_3)x_7 \leq (u_3 - m_3)V_{\text{old}} \\
& \quad l_4 x_1 + l_4 x_2 + l_4 x_3 + l_4 x_4 + l_4 x_5 + l_4 x_6 + (m_4 - l_4)x_7 \leq (m_4 - l_4)V_{\text{old}} \\
& \quad l_5 x_1 + l_5 x_2 + l_5 x_3 + l_5 x_4 + (l_5 - d_5)x_5 + l_5 x_6 \leq 0 \\
& \quad -x_5 \leq -0.5 \\
& \quad x_7 \leq V_{\text{old}} \\
\end{align*}
\]

\[x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0.\]
In this form, the replenishment problem is a standard Linear Programming (LP) problem, that can be solved by using for example the Simplex Algorithm. The Simplex Algorithm is described in every textbook on Operations Research (see e.g. Sierksma (1996)). An optimal solution of the problem is a set of values for the decision variables $x_1, x_2, \ldots, x_7$, such that the costs of replenishment are minimal while all restrictions are fulfilled.

### 8.2.4 A starting point for the Simplex Algorithm

Each time additions need to be computed, the LP model may differ from previous models because of different results of the analysis $(m_1, \ldots, m_5)$ and a different volume of the bath $(V_{\text{old}})$. Hence, assuming that the lower and upper limits and the concentrations of the additions remain unchanged, the model has six parameters.

If the model is solved by means of the Simplex Algorithm, a starting point (a so-called basic feasible solution) is needed. Generating a basic feasible solution can be done for example by means of the Big M method, or by the first phase of the Two Phase Simplex Method (see e.g. Sierksma (1996)). However, some computing efficiency could be gained if we had a starting point for the Simplex Algorithm that is independent of the measurements. We could just plug in our ‘standard’ basic feasible solution, and start the Simplex Algorithm right away, whatever the measurements are.

Such a measurement independent basic feasible solution exists if it is possible to create a new bath out of the separate components. This becomes clear if we realize that this corresponds to draining off the whole bath, i.e. setting $x_7$ equal to $V_{\text{old}}$, and create a completely new bath that satisfies the requirements. If such a solution exists, it is independent of the measurements since we started with an empty bath. In mathematical terms, this means choosing $x_7 = V_{\text{old}}$. The objective function can then be written as

$$\text{minimize } c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_4 + c_5 x_5 + c_6 x_6,$$

where we suppressed the term $c_7 V_{\text{old}}$, since this is a constant and therefore does not influence the optimal solution of the problem. The set of constraints reduces to
8.2. COMPUTING ADDITIONS TO THE TIN-PLATING BATH

\[ c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_4 + c_5 x_5 + c_6 x_6 \]
\[ - x_1 - x_2 - x_3 - x_4 - x_5 - x_6 \leq -b_{\text{min}} \]
\[ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \leq b_{\text{max}} \]
\[ (l_1 - d_1) x_1 + (l_1 - d_1) x_2 + l_1 x_3 + l_1 x_4 + l_1 x_5 + l_1 x_6 \leq 0 \]
\[ (d_1 - u_1) x_1 + (d_1 - u_1) x_2 - u_1 x_3 - u_1 x_4 - u_1 x_5 - u_1 x_6 \leq 0 \]
\[ l_2 x_1 + l_2 x_2 + l_2 x_3 + l_2 x_4 + l_2 x_5 + l_2 x_6 \leq 0 \]
\[ - u_2 x_1 + (d_2 - u_2) x_2 - u_2 x_3 - u_2 x_4 - u_2 x_5 - u_2 x_6 \leq 0 \]
\[ l_3 x_1 + l_3 x_2 + (l_3 - d_3) x_3 + l_3 x_4 + l_3 x_5 + l_3 x_6 \leq 0 \]
\[ - u_3 x_1 - u_3 x_2 + (d_3 - u_3) x_3 - u_3 x_4 - u_3 x_5 - u_3 x_6 \leq 0 \]
\[ l_4 x_1 + l_4 x_2 + l_4 x_3 + (l_4 - d_4) x_4 + l_4 x_5 + l_4 x_6 \leq 0 \]
\[ l_5 x_1 + l_5 x_2 + l_5 x_3 + l_5 x_4 + (l_5 - d_5) x_5 + l_5 x_6 \leq 0 \]
\[ - x_5 \leq -0.5 \]
\[ x_1, x_2, x_3, x_4, x_5, x_6 \geq 0, \]

Solving this model provides a ‘standard’ basic feasible solution to the replenishment problem that is independent of the measurements. This solution can be interpreted as the cheapest way to compose an admissible new tin-plating bath out of its separate components.

Note that the inclusion of the trivial constraint \( x_7 \leq V_{\text{old}} \) in the original model (which is not relevant for the optimal solution) makes such a solution a basic feasible solution.

8.2.5 Comparison of results

In this section we compare some solutions arrived at through the old method with the corresponding solutions of the LP model. To this end, we use four real analyses that were handed to us by a responsible technical engineer, whom we asked for a few measurements that are typical for the process. For these analyses, we will compare the output of the method currently used with the LP solution. The method currently used basically computes the additions one by one. In Subsection 8.2.1, we discussed why this approach may result in additions that are too large. Solving the LP model was done with the computer program PCProg. The size of the model is such that computing time is negligible (less than a second on a 486 personal computer).
CHAPTER 8. PHILIPS CASE

The analysis of July 10, 1995

In the analysis of the tenth of July, the concentration of acid fell between its upper and lower limits, but the concentrations of Sn\(^{2+}\) solution, Pb\(^{2+}\) solution, and brightener were too low. The bath volume was 470 liters, only 20 liters over the minimal bath volume, so that a considerable addition was possible before the bath would overflow.

Table 8.2: Comparison of replenishing strategies for July 10, 1995.

<table>
<thead>
<tr>
<th>component</th>
<th>old method</th>
<th>old method*</th>
<th>LP solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>acid</td>
<td>114.6 l</td>
<td>71.4 l</td>
<td>59.16 l</td>
</tr>
<tr>
<td>Sn(^{2+}) solution</td>
<td>135.0 l</td>
<td>84.1 l</td>
<td>74.99 l</td>
</tr>
<tr>
<td>Pb(^{2+}) solution</td>
<td>26.9 l</td>
<td>16.8 l</td>
<td>15.36 l</td>
</tr>
<tr>
<td>brightener</td>
<td>27.3 l</td>
<td>17.0 l</td>
<td>16.00 l</td>
</tr>
<tr>
<td>formalin</td>
<td>3.5 l</td>
<td>2.2 l</td>
<td>1.67 l</td>
</tr>
<tr>
<td>demineralized water</td>
<td>0.0 l</td>
<td>0.0 l</td>
<td>0.00 l</td>
</tr>
<tr>
<td>drain off</td>
<td>0.0 l</td>
<td>177.3 l</td>
<td>37.18 l</td>
</tr>
<tr>
<td>new bath volume</td>
<td>777.3 l</td>
<td>484.2 l</td>
<td>600.00 l</td>
</tr>
</tbody>
</table>

**total costs:**

Dfl. 8317.48 Dfl. 6853.01 Dfl. 3088.33

*draining off before making additions allowed

In Table 8.2, the results of the old method and the LP solution are tabulated. Note that there are two columns for the solution of the old method. In the first column, indicated by ‘results of old method’ the rough results of the old method can be found. From this we see that, since the maximum bath volume \(b_{\text{max}} = 600\) liters, the bath would overflow with 177.3 liters. This suggestion is therefore not admissible. Furthermore, since this method does not take draining off into account, the additions are computed on the basis of a high volume bath. Apart from any computing inefficiencies, this is another reason why the additions are so high.

The second column presents the results of running the program with the same initial concentration measurements, but now some tin-plating liquid is allowed to be drained off first. Usually the operators drained off an amount equal to the surplus of the first suggestion (in this case 177.3 liters). When draining off before making additions is allowed, both the amounts of chemicals used and the total costs are reduced. Total costs decrease with approximately 1500 Dutch guilders. However, an even greater cost
reduction is possible if we use the LP-solution, which is tabulated in the last column of Table 8.2. The main part of the cost reduction stems from finding a combination of the decision variables so that the number of liters that is to be drained off is kept at a minimum. As a result, the new bath volume is maximal, while the new bath volume of the old method is close to its minimal value of $b_{\text{min}} = 450$ liters.

The analysis of August 29, 1995

The concentration of Sn$^{2+}$ solution in the tin-plating bath on the 29th of August was a little too high. All other concentrations were acceptable. The bath volume was 575 liters, so that only small additions would be allowed. Fortunately, the bath had to be thinned only a little bit. In Table 8.3 we present the results of the two strategies. Note that in this table only one column is presented for the output of the old method since draining off does not appear to be necessary.

Table 8.3: Comparison of replenishing strategies for August 29, 1995.

<table>
<thead>
<tr>
<th>component</th>
<th>old method</th>
<th>LP solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>acid</td>
<td>0.0 l</td>
<td>0.00 liters</td>
</tr>
<tr>
<td>Sn$^{2+}$ solution</td>
<td>0.0 l</td>
<td>0.00 liters</td>
</tr>
<tr>
<td>Pb$^{2+}$ solution</td>
<td>2.2 l</td>
<td>0.00 liters</td>
</tr>
<tr>
<td>brightener</td>
<td>4.4 l</td>
<td>0.00 liters</td>
</tr>
<tr>
<td>formalin</td>
<td>0.4 l</td>
<td>0.50 liters</td>
</tr>
<tr>
<td>demineralized water</td>
<td>6.0 l</td>
<td>2.90 liters</td>
</tr>
<tr>
<td>drain off</td>
<td>0.0 l</td>
<td>0.00 liters</td>
</tr>
<tr>
<td>new bath volume</td>
<td>588.0 l</td>
<td>578.40 liters</td>
</tr>
<tr>
<td>total costs:</td>
<td>Dfl. 111.15</td>
<td>Dfl. 0.69</td>
</tr>
</tbody>
</table>

In an absolute sense, the differences between the methods are less spectacular than in the previous case, but the cost reduction is enormous if expressed in percentages. The total costs of the LP solution are only 0.6% of the total costs of the old method. The LP solution adds the obligatory 0.5 liters of formalin, and some demineralized water. The old method suggests to add more water, together with Pb$^{2+}$ solution and brightener. The last two additions explain the main part of the cost difference between the
methods. Note that the old method apparently has no lower limit on the addition of formalin.

The analysis of September 27, 1995

On the 27th of September, the tin-plating bath met all but one of the requirements: the concentration $\text{Sn}^{2+}$ solution was too low. The concentration of $\text{Pb}^{2+}$ solution was found exactly at its lower limit, so that with any addition of another component, addition of $\text{Pb}^{2+}$ solution would become necessary. The bath volume was 475 liters, so that quite large additions were possible. In Table 8.4 the results of the two methods can be found. Since the output of the old method initially results in an overflowing bath, again a second column is presented in which there is allowance for draining off the surplus of the first solution before computing additions. The last column contains the solution of the LP model.

Table 8.4: Comparison of replenishing strategies for Sept. 27, 1995.

<table>
<thead>
<tr>
<th>Component</th>
<th>Old method</th>
<th>Old method*</th>
<th>LP solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>acid</td>
<td>67.2 l</td>
<td>61.0 l</td>
<td>32.16 l</td>
</tr>
<tr>
<td>$\text{Sn}^{2+}$ solution</td>
<td>80.0 l</td>
<td>72.5 l</td>
<td>43.82 l</td>
</tr>
<tr>
<td>$\text{Pb}^{2+}$ solution</td>
<td>15.0 l</td>
<td>13.6 l</td>
<td>6.76 l</td>
</tr>
<tr>
<td>brightener</td>
<td>4.9 l</td>
<td>4.5 l</td>
<td>0.00 l</td>
</tr>
<tr>
<td>formalin</td>
<td>2.0 l</td>
<td>1.8 l</td>
<td>0.84 l</td>
</tr>
<tr>
<td>demineralized water</td>
<td>0.0 l</td>
<td>0.0 l</td>
<td>0.00 l</td>
</tr>
<tr>
<td>drain off</td>
<td>0.0 l</td>
<td>44.1 l</td>
<td>0.00 l</td>
</tr>
<tr>
<td>new bath volume</td>
<td>644.1 l</td>
<td>584.2 l</td>
<td>558.60 l</td>
</tr>
<tr>
<td>total costs:</td>
<td>Dfl. 3149.17</td>
<td>Dfl. 2961.59</td>
<td>Dfl. 1032.04</td>
</tr>
</tbody>
</table>

*draining off before making additions allowed

The LP solution is approximately 2000 Dutch guilders cheaper than the solution of the old method. The cost reduction is caused by a reduction in the use of chemicals. The effect of this is twofold; not only do we have a reduction in the costs of the additions, it also becomes unnecessary to drain off tin-plating liquid, which is very expensive.
The analysis of October 19, 1995

The analysis of the tin-plating bath of the final example indicated that the concentration of Sn\(^{2+}\) solution was too low, while the concentration Pb\(^{2+}\) solution was too high. The other concentrations fell within their limits. The bath volume was 575 liters, so that there was not much room for making additions. According to Table 8.5 both methods do not appear to be able to find additions such that draining off can be avoided.

<table>
<thead>
<tr>
<th>component</th>
<th>old method</th>
<th>old method*</th>
<th>LP solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>acid</td>
<td>24.6 l</td>
<td>23.1 l</td>
<td>12.99 l</td>
</tr>
<tr>
<td>Sn(^{2+}) solution</td>
<td>33.9 l</td>
<td>31.9 l</td>
<td>20.44 l</td>
</tr>
<tr>
<td>Pb(^{2+}) solution</td>
<td>0.0 l</td>
<td>0.0 l</td>
<td>0.00 l</td>
</tr>
<tr>
<td>brightener</td>
<td>0.0 l</td>
<td>0.0 l</td>
<td>0.00 l</td>
</tr>
<tr>
<td>formalin</td>
<td>0.0 l</td>
<td>0.0 l</td>
<td>0.50 l</td>
</tr>
<tr>
<td>demineralized water</td>
<td>0.9 l</td>
<td>0.9 l</td>
<td>2.69 l</td>
</tr>
<tr>
<td>drain off</td>
<td>0.0 l</td>
<td>36.4 l</td>
<td>11.62 l</td>
</tr>
<tr>
<td>new bath volume</td>
<td>636.4 l</td>
<td>596.5 l</td>
<td>600.00 l</td>
</tr>
<tr>
<td><strong>total costs:</strong></td>
<td>Dfl. 1675.07</td>
<td>Dfl. 1628.83</td>
<td>Dfl. 743.03</td>
</tr>
</tbody>
</table>

*draining off before making additions allowed

In this case too, the LP solution results in a considerable cost reduction of replenishment.

8.2.6 Conclusions

In this section, we discussed an alternative way of computing additions for replenishing a chemical bath that needs regular adjustment. To illustrate what cost reductions were possible in practice, we compared the replenishments that were computed by the method currently used to the corresponding LP solutions, for four typical conditions of the bath. The total costs of the method currently used were 13,252.87 guilders if no draining off before adding was allowed, and 11,558.58 guilders if draining off was allowed. In contrast, the total costs of the LP solution were 4,864.09 guilders for these four days. This means that a cost reduction of 63% and 58%, respectively is possible! Philips Stadskanaal has several tin-plating baths, which are
replenished on a regular basis.

The methodology described in this section is easily adjusted to accommodate similar replenishment problems. When all the components are separately available, the corresponding constraints follow immediately. However, if some of the components can only be added by adding a mix of components, as was the case with the Sn$^{2+}$ solution, the constraints can easily be adapted to deal with this situation. If some of the components are separately available, but also in some pre-mixed form (which may be cheaper), the model can consider all of these as possible additions, and choose the cheapest combination.

Depending on the number of components, the LP-model is typically small. With the aid of modern computer equipment such a model is solvable within a second. However, it is possible to gain a little computation time if a standard starting solution for the Simplex Algorithm is used.

We would like to emphasize that this LP-model is only a tool for computing the cheapest way to create a bath that satisfies certain predefined limits. The cost reductions stem only from a more efficient method of computing the additions, not from a radical change in the replenishment strategy.

However, as the quality improvement project progressed, we obtained further insight in the process by studying measurements that were taken from the production process. Based on the observed data, our proposal was essentially to part with the daily replenishments. The arguments for this piece of advice are described in the remainder of this chapter. The PAT decided to experiment with this drastic change in the replenishment strategy, and the results were surprisingly good. As a result, the LP method, so far, has not been used on the production floor of the department where the SMID’s are produced. However, in departments where other types of diodes are manufactured the method has been successfully implemented.

8.3 Adjusting the tin-plating bath

In Chapter 2 we referred to Nolan and Provost (1990)) for the following definition of a process: “a set of causes and conditions that repeatedly come together to transform inputs into outcomes”, see Figure 8.3.

For the process under consideration, one of the inputs is the composition of the tin-plating bath. Relevant outputs are the thickness and the composition of the tin/lead layer. A lot of effort is put into maintaining a stable tin-plating bath with the objective of reducing variation in the
8.3. ADJUSTING THE TIN-PLATING BATH

outputs of the process. In this section, we will investigate if such efforts can be justified by studying process data. In Subsection 8.3.1, we describe the data that is available. In Subsection 8.3.2, we investigate whether inputs data has a relation with data on the output side of the process. The measurement error of the observations of the process is discussed in Subsection 8.3.3. In Subsection 8.3.4, it is argued why the presumed relationships between inputs and outputs are not reflected in the data. The conclusion of this investigation and its implications for running the process are discussed in Subsection 8.3.5.

8.3.1 The data

Two kinds of measurements on the tin-plating process are available. In addition to measurements of the inputs of the tin-plating process, we have output measurements of the products at the end of the tin-plating process at our disposal.

The input measurements consist of a chemical analysis of the tin-plating bath. By means of titration the concentrations of acid, Sn$^{2+}$ and Pb$^{2+}$ are measured. The concentration of brightener can only be roughly determined, and there is no measurement device available for establishing the concentration of formalin. The analysis is performed once a day by the operators. The results form the input for a computer program that computes the necessary additions in order to obtain a bath that is fit for applying a good tin/lead layer.

The limits that are used by this program are *process limits*: during production, the concentration measurements are not supposed to exceed these limits. Hence, a replenishment is only prescribed if the observed concentrations are out of specifications. Consequently, the process was functioning outside its process limits half of the time according to the measurements. Fortunately but surprisingly, the quality of the tin/lead layers did not deteriorate in such situations. This gives good reason to doubt the strategy that was used to control the composition of the bath.

Moreover, the value of the process limits is hard to explain. The supplier of the tin-plating chemicals prescribes a certain ideal composition of
the tin-plating bath. However, in the course of time these limits were adjusted, partly based on experience of the operators, and partly based on superstition.

Furthermore, compared to the variation in the measurements, the limits do not seem to be equally tight for all of the components. This is illustrated by Figures 8.4 and 8.5, where respectively the concentrations of Sn\(^{2+}\) solution and Pb\(^{2+}\) solution are depicted. The graphs reflect the results from consecutive analyses over a three weeks period.

![Figure 8.4: The behavior of Sn\(^{2+}\) solution in time.](image)

From these figures we see that the process limits for the Sn\(^{2+}\) solution are tighter compared to the variation in the measurements than the limits for Pb\(^{2+}\) solution. It is to be expected that the concentration of Sn\(^{2+}\) solution needs to be adjusted more often than the concentration of Pb\(^{2+}\) solution.

Another peculiarity of Figures 8.4 and 8.5 is that the phenomenon of bath exhaustion is not noticeable from these figures. Instead of concentrations jumping up and down more or less randomly, we would have expected a trend or some other form of deterministic behavior (possibly depending on replenishments).

The available output measurements consist of samples of size 10 that are taken from every batch of diodes. A batch may contain up to 31000 diodes and about 60 batches are processed each day in three shifts. From each
of the sampled diodes, the thickness and the composition of the tin/lead layer are measured. The frequency of sampling and the sample size were set some years ago, and none of the people involved remembered the exact reasons for this sampling strategy. The operators spent considerable time taking these measurements. During the PAT meetings, we first decided to take a sample every other batch, and later on to reduce the sample size to five as well.

A sample size of 10 observations suffices to obtain an impression of the performance of the process. This makes the data fit for process monitoring. However, the data were used for acceptance sampling. On the basis of a sample of size 10 it was decided to reject or accept a batch of 31000 diodes! The following decision rule was used: if the thickness of the tin/lead layer of one or more of the sampled diodes did not exceed 3μm, or if the sample mean was smaller than 4.85μm, it was decided to tin-plate the whole batch again. Otherwise, the batch was accepted and passed on to the inspection department where the following production step takes place. This sampling plan is not sufficient to ensure a high quality level of the outgoing batches.

For example, assuming independence and normality of the observations and a (realistic) standard deviation equal to 4, a batch with 5% of the tin/lead layers smaller than 1μm passes this test with a probability of about 25%.

Figure 8.5: The behavior of Pb^{2+} solution in time.
Indeed, if one of these tests rejects the batch, then there are good reasons to suspect that something is wrong. However, to ensure an outgoing quality level of only a few defective parts per million, the sample size must be increased drastically.

In the next subsection, we will investigate whether there are some sensible relationships between inputs and outputs visible in the data.

8.3.2 The relation between input and output measurements

In this subsection, we consider a data set of 106 observations of bath analyses and corresponding product measurements. The observations were successively taken on a daily basis in the last five months of 1995.

The bath analyses are approximately equally spaced, and have values for the concentration acid, $\text{Sn}^{2+}$ solution, $\text{Pb}^{2+}$ solution and brightener. Recall from the previous subsection that the concentration of brightener can only be roughly determined, whereas the concentration of formalin is not measured at all. The volume of the bath, on the other hand, is available. The output measurements are taken from the last produced batch before analyzing and replenishing the bath. A sample of ten diodes is randomly drawn from the batch and the average thickness and the average composition of the tin/lead layers are recorded.

After investigating several (lagged) relationships between inputs and outputs, suggested by the chemical reactions taking place in the bath, we reached the conclusion that none of the presumed relationships was reflected in the data. To be more precise, we were not able to find any significant relation between input and output measurements that had a practical importance.

Note that this might have been expected from Figures 8.4 and 8.5, where we observed that the bath was frequently operating outside its process limits, while the tin/lead layers were satisfactory.

As an illustration of the lack of correlation, we depicted in Figure 8.6 a scatter plot of the ratio of the $\text{Sn}^{2+} / \text{Pb}^{2+}$ concentrations in the bath, against the $\text{Sn}^{2+} / \text{Pb}^{2+}$ ratios of the product measurements. Contrary to everyone’s expectations, no strong correlation is present.

From the arguments above, we realized that attempting to control the quality characteristics of tin/lead layers on the basis of the bath measurements is asking for trouble. This conclusion meant a breakthrough in the attempts of the PAT to improve the quality of the tin/lead layers, since it opened the way for other means of controlling the output of the process.
8.3. ADJUSTING THE TIN-PLATING BATH

However, before developing a new control mechanism, it had to be understood what caused the lack of correlation between inputs and outputs. To this end the accuracy of the measurements was evaluated first. The results are summarized in the following subsection. Secondly, chemical experts were invited to explain why graphs similar to Figure 8.6 did not show some expected relationship.

![Figure 8.6](image)

**Figure 8.6:** The relation between Sn\(^{2+}\)/Pb\(^{2+}\) in the bath and on the product.

8.3.3 Measurement errors

Based on the methodology described in Does, Roes and Trip (1999), or Banens et al. (1994), an R&R (Repeatability & Reproducibility) study for the product measurements had already been performed. It turned out that the ‘gage R&R’ was 1.120 for the thickness of the tin/lead layers, while the ‘gage R&R’ was 2.346 for the concentration measurements. The ‘gage R&R’ is computed as 5.15\times\text{(total measurement variance)}, and under the assumption of normally distributed observations it is the width of a 99% confidence interval. To judge the accuracy of the measurement device, the ‘gage R&R’ is compared to the width of the specification limits. In our
case, the product measurements were sufficiently accurate.

Considering the bath measurements, we reached another conclusion. To determine the concentrations of the bath components, operators were asked to take exactly 1 ml of bath liquid using a pipette, whereafter an automated titration device determined how much (in mol) of each component was present in the liquid. The results of these measurements were imported into an online computer, which divided these numbers by 1 ml and presented the outcomes as concentration measurements. Clearly, the accuracy of this procedure is largely determined by the ability of the operators to pipette exactly 1 ml of bath liquid. To investigate this, operators were asked to pipette at least five times 1 ml of demineralized water. To determine the pipetted quantity, it was weighted on a very accurate pair of scales. The results for 21 operators are depicted in Figure 8.7.

![Box plots of pipetting results.](image)

In the Box plots of Figure 8.7 the median of the observations per operator is indicated by a white line in the box. The length of the boxes indicates the interquartile range, and the connected hooks (if present) show the minimum and the maximum values. Observations falling outside the range of $1.5 \times$ (interquartile range) are considered as outliers, and are plotted as disconnected dots. The width of the boxes is proportional to the square root
of the number of observations per operator. The target value of 1 ml is indicated by a solid line.

Figure 8.7 shows that there is considerable variation between operators. Often, the target value of 1 ml is not even in the range of the observations. Some operators are able to take the amount of water with high precision, but in most of these cases the level is off-target. In other cases the range is quite large, which is not always the result of a larger number of observations.

These observations were taken in a laboratory at the end of an operator course. It is therefore to be expected that the results on the work floor will be even less accurate. Furthermore, Figure 8.7 only shows the variation due to pipetting. There is also additional variation, for example variation due to measurement error of the titration device. Hence, it is to be expected that the total variation in the analysis measurements is even higher than the variation shown in Figure 8.7. However, if we use the results of this experiment, and assume that not water but bath liquid was pipetted, and that the true concentration values were all exactly on target (i.e. in the middle of specification limits), then we are able to compute what the effect of the pipetting inaccuracy would have been on the analysis results. For the concentration of Sn$^{2+}$, this is depicted in Figure 8.8.

![Box plots of Sn$^{2+}$ measurement inaccuracies.](image)
CHAPTER 8. PHILIPS CASE

From Figure 8.8 it may be concluded that the measurement error due to pipetting is unacceptably high. The analysis results of a bath, the true concentration values of which are exactly in the middle of the specification limits (an ideal bath), can lead to the absurd conclusion that the bath is not fit for production. The analyses tell us more about the large measurement error than about the contents of the bath. The conclusion is that controlling on the basis of analysis results essentially comes down to controlling on the basis of measurement error. Such control actions generally increase the variation in the outcomes of the process. This type of overacting was called tampering with the process by Deming.

8.3.4 Process mechanics

In our search for an explanation for the lack of correlation between observed process inputs and process outputs, we also discovered that the assumptions concerning the manner in which the inputs affect the outputs were not entirely correct. Since such assumptions form the basis for any control strategy, this gives another explanation for the malfunctioning of the current replenishment strategy.

Recall from Section 8.1.4 that the tin-plating bath consists of five components: acid, an Sn$^{2+}$ solution, a Pb$^{2+}$ solution, brightener, and formalin. The current strategy calls for action if one or more of the concentrations of Sn$^{2+}$ solution, Pb$^{2+}$ solution, or acid is out of range. These three parameters are thought to have a great impact on the composition and the thickness of the tin/lead layers. The concentration of brightener is only very roughly determined, whereas determining the concentration of formalin is not possible at all. Hence, controlling basically takes place on the first three parameters. However, in brainstorming sessions together with chemical engineers, it turned out that the thickness and the composition of the layer were mainly determined by the two other parameters. Brightener not only takes care of a shiny tin/lead layer, it also influences the thickness of the layer. The concentration of formalin is an important determinant of the composition of the tin/lead layers.

The other components are not unimportant: the acid takes care of a good conductivity of the bath, while the concentrations of the Sn$^{2+}$ and Pb$^{2+}$ solutions must be large enough to ensure that the chemical reactions as described in Section 8.1.4 can take place. However, none of these does directly influence the characteristics of the tin/lead layer on the product.

Furthermore, it was assumed that due to the primary chemical reactions in production, the concentrations of Sn$^{2+}$ and Pb$^{2+}$ solution would
change. If we take another look at the chemical reactions, this does not make sense. For every Sn$^{2+}$ or Pb$^{2+}$ ion that precipitates on a diode, another one dissolves from the anode (see Figure 8.2). So, apart from secondary chemical reactions, there is no reason to assume that the concentration of Sn$^{2+}$ solution or the concentration of Pb$^{2+}$ solution will change with the volume of production.

### 8.3.5 Conclusion

From the discussion in this chapter, it is easy to deduce why the existing control strategy did not lead to the desired results. Firstly, control actions took place on the basis of relatively unimportant process parameters. And secondly, the measurements of these parameters are highly inaccurate, so that the control actions are not based on process information, but rather on ‘noise’. As a result, the effort that was put into stabilizing the bath, caused the bath to destabilize. This type of overacting is called tampering, and is a well-known phenomenon in the field of statistical quality control. It is one of the main arguments for using SPC techniques. In the next section we will see that it is possible to obtain a more stable process by simply not reacting on input measurements. A control strategy is proposed that is based on leaving the process alone, except in situations where there is statistical evidence of the presence of an additional, abnormal source of variation.

### 8.4 A new control strategy

In the previous section we evaluated the control strategy that was used up until now. We reached the conclusion that this strategy, due to tampering, has a destabilizing effect on the bath. Furthermore, we argued that from the chemical reactions discussed in Section 8.1.4, it is not clear why replenishments are necessary at all.

Therefore, we experimented with producing without replenishments (except for formalin for reasons of quick evaporation, and brightener). Not surprisingly, the variation in the measurements of the tin/lead layers decreased. During the period of not replenishing, we closely monitored the product measurements to see whether this strategy had the expected effect. Details can be found in Sections 8.4.1 and 8.4.2. After four weeks of applying the new strategy, it was observed that one of the quality parameters, the composition of the tin/lead layer, was drifting away. From a reliable analysis of the bath, performed by the laboratory, the level of acid and
Sn$^{2+}$ solution turned out to be too low for a stable process. Very likely, this was due to secondary chemical reactions in the bath.

Based on the experiment described above, we dropped the daily analysis by the operators and devised a new replenishment strategy, based on not replenishing except in situations where there is statistical evidence of something abnormal affecting the process (and except for daily formalin and brightener additions). However, the experiment showed that four weeks without replenishing was too long.

Once a week the bath is filtered to clean the tin-plating liquid from precipitations, caused by secondary reactions. This was a natural moment for a weekly reliable bath analysis, performed by a laboratory employee. On the basis of this analysis, appropriate replenishments can be determined with the aid of the LP model, discussed in section 8.2, to make sure that the concentrations are within their limits.

During the week, the product measurements are monitored with the aid of a control chart, to check whether there is evidence of an additional, abnormal source of variation. How this is done will be discussed in the following subsections.

### 8.4.1 Monitoring the process

In Figures 8.9 and 8.10 the product measurements of the period March 4, 1996 through March 31, 1996 are depicted. This was the period in which it was decided to experiment with producing without replenishing.

Figure 8.9 shows the successive means of thickness measurements, and Figure 8.10 shows the successive means of Sn$^{2+}$ /Pb$^{2+}$ ratios. Remember that a sample of size 10 is taken from every batch of approximately 31,000 diodes.

Despite the agreement not to make additions, it can be suspected from both figures that around observation 300 something unusual has happened to the process. Both the layer thickness and composition values seem to be affected by some special cause of variation.

Figure 8.10 also shows some other interesting behavior. The mean composition values are, on average, slowly decreasing. This trend is only interrupted by the same sudden jump upwards around observation 300. We will discuss these phenomena in more detail in Subsection 8.4.4.

Figure 8.11 shows a scatter plot of layer thickness and composition. No strong relationship seems present. The correlation coefficient between thickness and composition measurements can be computed as 0.21. We therefore decided to monitor the process with two univariate control charts.
8.4. A NEW CONTROL STRATEGY

Figure 8.9: Successive mean thickness observations.

Figure 8.10: Successive mean Sn\(^{2+}\)/Pb\(^{2+}\) ratio observations.
In Chapter 1, it was discussed that the application of standard control charts hinges on two important assumptions: normality and independence of successive observations. Since serial correlation distorts various tools for checking normality of the data (such as normal probability plots), the correlation structure of the data is explored first, and the exploration of normality is deferred to Subsection 8.4.3.

A useful tool to explore the correlation structure of the data is the sample autocorrelation function. Figures 8.12 and 8.13 depict the sample autocorrelation functions for mean layer thickness, and mean layer composition, respectively.

It can be shown (see e.g. Anderson (1971)) that a fixed number of sample autocorrelation coefficients of white noise are asymptotically normally and independently distributed with zero means and standard deviations equal to $1/\sqrt{N}$, where $N$ is the number of observations, in our case equal to 644. This can be used to judge the sample autocorrelation coefficients.

The dashed lines in Figures 8.12 and 8.13 are drawn at $1.96 \times 1/\sqrt{N}$ as an approximate 95% confidence interval for individual sample autocor-

**Figure 8.11:** Scatter plot of mean Sn$^{2+}$/Pb$^{2+}$ ratios and mean layer thickness.
Figure 8.12: Autocorrelation function of mean thickness values.

Figure 8.13: Autocorrelation function of mean composition values.
relations with expectation zero. Note that since we are plotting a number of autocorrelation coefficients at once, we expect to find, on average, one out of twenty outside these limits if the observations are uncorrelated. In addition, if the covariance between successive observations is nonzero, the sample autocorrelations are also correlated. These phenomena may seriously distort the interpretation of a sample autocorrelation function, and conclusions based upon this graph must be drawn with care.

Notwithstanding the above remarks, the conclusion from Figures 8.12 and 8.13 is that both layer thickness and layer composition are highly autocorrelated. Moreover, both autocorrelation functions suggest nonstationarity. When constructing control charts, we must take these findings into account. In the next section we will discuss why this is necessary.

8.4.2 Ignoring the serial correlation

First we ignore the serial correlation and set up a control chart for the mean of mean layer composition in the usual way; that is, assuming independence and normality of the observations.

Let us denote the observed mean composition value on time \( t \) by \( \bar{y}_t \). Remember that means are taken over samples of size \( n = 10 \). Trial control limits with tail probabilities of \( \frac{1}{2}\alpha \) are determined as (see Does and Schriever (1992))

\[
\text{LCL} = \bar{y} + \sqrt{\frac{N - 1}{Nn}} \ t_{1 - \frac{1}{2}\alpha}(N(n - 1)) \hat{s},
\]

and

\[
\text{UCL} = \bar{y} + \sqrt{\frac{N - 1}{Nn}} \ t_{1 - \frac{1}{2}\alpha}(N(n - 1)) \hat{s},
\]

where LCL and UCL stand for Lower Control Limit and Upper Control Limit, respectively, and \( \bar{y} \) is the overall mean of \( N = 644 \) observations of \( y_t \) (\( t = 1, \cdots, 644 \)). Furthermore, \( t_{1 - \frac{1}{2}\alpha}(N(n - 1)) \) is the \( \frac{1}{2}\alpha \)-th percentile of the \( t^2 \) distribution with \( N(n - 1) \) degrees of freedom, and \( \hat{s} \) is defined as

\[
\hat{s} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} s_i^2},
\]
allow for serial correlation, is not a good idea in this case. Therefore, we formed measurements are compared to control limits which are adjusted to stabilize the mean, a lot of out-of-control signals will be generated by ARIMA($p; d; q$) models. A feature of an IMA(1,1) process is that its level is changing constantly. It is therefore to be expected that, in testing for stability of the mean, a lot of out-of-control signals will be generated by such a data set.

Hence, the data must be monitored using a control chart that takes serial correlation into account. Using a modified control chart, where untransformed measurements are compared to control limits which are adjusted to allow for serial correlation, is not a good idea in this case. Therefore, we revert to residuals charts to monitor the process of tin-plating diodes.

In the next subsection, an appropriate time series model is selected and fitted to the data.
Finding an appropriate model for the data

The autocorrelation function of Figure 8.13 shows considerable autocorrelation which seems to decrease slowly and linearly with the length of the lags, an indication for the presence of a unit root. Box and Jenkins (1976) explain why this is so.

Let a sequence of random variables \( \{Z_t\} \) follow a stationary ARMA\((p,q)\) model. Assume that \( \text{E}(Z_t) = 0 \). Then we have that \( Z_t \) can be written as

\[
Z_t = \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_q \epsilon_{t-q},
\]

where \( \{\epsilon_t\} \) is a white noise process. When \( B \), the backward shift operator is used, (8.1) can be written as

\[
(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) Z_t = (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q) \epsilon_t
\]

or

\[
\phi(B)Z_t = \theta(B)\epsilon_t,
\]

where \( \phi(\cdot) \) and \( \theta(\cdot) \) are polynomials of degree \( p \) and \( q \), respectively.

Multiplying both sides of Equation (8.1) by \( Z_{t-k} \) and taking expectations, we find that

\[
\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p} \quad \text{for } k \geq q + 1,
\]

where \( \gamma_k \) is the covariance between \( Z_t \) and \( Z_{t-k} \). For \( k \geq q+1 \) the covariance between \( Z_t \) and \( \epsilon_{t-k} \) is zero. Dividing both sides of Equation (8.2) by \( \gamma_0 \), the variance of \( Z_t \), this results in

\[
\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p} \quad \text{for } k \geq q + 1,
\]

where \( \rho_k \) is the autocorrelation coefficient of \( Z_t \) and \( Z_{t-k} \). And hence, the autocorrelation function satisfies the difference equation

\[
\phi(B)\rho_k = 0 \quad \text{for } k \geq q + 1.
\]

Furthermore, if it is assumed that \( \phi(B) = \prod_{i=1}^p (1 - G_i B) \), with \( G_1, \cdots, G_p \) distinct, so that \( 1/G_1, \cdots, 1/G_p \) are distinct roots of \( \phi(\cdot) \), then \( \rho_k \) is of the form
8.4. A NEW CONTROL STRATEGY

\[ \rho_k = A_1G_1^k + A_2G_2^k + \cdots + A_pG_p^k \quad \text{for} \ k \geq q + 1. \]

If one or more of the roots of $\phi(\cdot)$ approaches one, say $G_1 = 1 - \nu$ with $\nu > 0$ small, then

\[ \rho_k \approx A_1(1 - k\nu), \]

and the autocorrelation function will not die out quickly. Instead, it will be linearly decreasing in $k$.

On the other hand, for a model to be stationary, it is required that the roots of $\phi(\cdot)$ must lie outside the unit circle, so that $G_1, \ldots, G_p$ must lie inside the unit circle. Therefore, if none of the roots of $\phi(\cdot)$ is close to one, $\rho_k$ will damp out quickly.

Hence, Figure 8.13 shows the typical behavior of a nonstationary process. The next step is to compute first differences of the mean composition values. The autocorrelation function of the new series will give insight in the remaining correlation structure. If taking first differences does not remove nonstationarity, we take first differences once more, and so on, until the series displays stationary behavior. The sample autocorrelation function of first differences of the mean composition values is depicted in Figure 8.15.

The autocorrelation function of Figure 8.15 shows a single spike at lag 1, indicating a Moving Average (MA) component of order 1 in the first differences. This can be seen from the theoretical autocorrelation function of a first-order MA process. For a sequence of random variables $\{Z'_t\}$, generated by an MA(1) model, i.e.

\[ Z'_t = \varepsilon_t - \theta \varepsilon_{t-1}, \]

with $\{\varepsilon_t\}$ white noise, we have for the first-order correlation coefficient

\[ \rho(1) = \frac{E(Z'_t Z'_{t-1})}{V(Z'_t)} = \frac{-\theta}{1 + \theta^2}, \]

whereas $\rho(2) = \rho(3) = \cdots = 0$. Furthermore, a first-order MA process can also be viewed as an infinite-order AutoRegressive (AR) process with coefficients that decrease exponentially in absolute value:

\[ Z'_t = \varepsilon_t - \theta \varepsilon_{t-1} = (1 - \theta B)\varepsilon_t \]

\[ \iff \quad (1 - \theta B)^{-1}Z'_t = \varepsilon_t \]

\[ \iff \quad Z'_t = \varepsilon_t - \theta Z'_{t-1} - \theta^2 Z'_{t-2} - \theta^3 Z'_{t-3} + \cdots, \]
so that the theoretical partial autocorrelation function (see for example Harvey (1993)) of a first-order MA process is exponentially declining in absolute value. So, if an MA(1) model is useful for modelling the first differences, we expect to see such behavior in the corresponding partial autocorrelation function. The observed partial autocorrelation function of first differences of the mean composition values of Figure 8.16 does indeed show such behavior.

Hence, based on the behavior of the (partial) autocorrelation functions of the data, we decide to select the IMA(1,1). The maximum likelihood estimate of the moving average coefficient was computed with the aid of SPLUS as \( \hat{\theta} = 0.811 \), with a standard error of 0.023. The model for the series of mean composition values then becomes

\[
\tilde{y}_t = \tilde{y}_{t-1} + \varepsilon_t - 0.811 \varepsilon_{t-1}.
\]  
(8.5)

When model (8.5) is fitted to the data, the sample autocorrelation function of the residuals, as depicted in Figure 8.17, may be used to check whether there is still some remaining serial correlation present in the residuals.
Figure 8.16: Partial autocorrelation function of differenced mean composition values.

Figure 8.17: Autocorrelation function of residuals.
A warning regarding the interpretation of the residual sample autocorrelation function must be made at this point. Granger and Newbold (1986) cite references from which it follows that the asymptotic standard deviations of sample autocorrelation coefficients of residuals may be smaller than $1/\sqrt{N}$. Hence, the sample autocorrelation coefficients of Figure 8.17 need careful interpretation. Nevertheless, the lines drawn in this figure at $1.96 \times 1/\sqrt{N}$ can provide a crude check for model adequacy. These limits are only just exceeded at lag 1 and 15.

However, the autocorrelation coefficients are small, especially when compared to the autocorrelation function of Figure 8.13, so that the correlation is considerably reduced.

Also, evaluation of the well-known (modified) Portmanteau test statistic at various lags does not lead to rejection of one of the hypotheses of zero correlations. Hence, for this moment, we proceed as if the residuals are uncorrelated.

In Subsection 8.4.1, we deferred testing for normality because of the presence of serial correlation. Since the residuals of model (8.5) behave more like an uncorrelated sequence than the original observations, its normal probability plot is more reliable for judging normality than a normal probability plot of the original observations. In Figure 8.18 a normal probability plot of the residuals of model (8.5) is shown. The normal probability plot does not indicate deviations of normality. Also more formal tests do not reject the hypothesis that the residuals are normally distributed. For example, a test based on observed kurtosis and skewness of the empirical distribution function as described by Harvey (1993), p. 45 results in a $p$-value of 0.330.

For the remainder of this chapter we assume that model (8.5) is the correct model for our data. The residuals of the fitted model behave more or less as a sequence of uncorrelated normally distributed variates. Hence, both conditions discussed on page 194 are fulfilled, and the standard control charts can be used to monitor the residuals. As we will see, not only the residuals contain information about changes in the process. It is wise to monitor the fitted values as well. In the next subsection we will discuss how the results of the data analysis can provide us with information about changes in the process.

### 8.4.4 The common cause chart

In the previous subsection, we modelled the mean composition observations using an IMA(1,1) model. The process generates observations with a
wandering level, also in cases where only common causes of variation are affecting the process. A graph of the level of the process now contains nontrivial information, viz. how the process is affected by common causes. Alwan and Roberts (1988) call such a graph a common cause graph, since it shows the variation due to causes of variation that are inherent in the process. If the process is not allowed to wander too far from a certain target value, some kind of active control must be applied to assure that product measurements do not fall outside specification limits. Let us consider the sequence of fitted values.

For \( \hat{y}_t \), the optimal one-step ahead forecast of \( \bar{y}_t \), we have

\[
\hat{y}_t = \mathbb{E} ( \bar{y}_t | \bar{y}_{t-1}, \bar{y}_{t-2}, \bar{y}_{t-3}, \ldots )
\]

\[
= (1 - \theta) \sum_{j=1}^{\infty} \theta^{j-1} \bar{y}_{t-j},
\]

so that the optimal forecast of \( \bar{y}_t \) is an Exponentially Weighted Moving Average (EWMA) of previous observations. In Chapter 4, we explained that the EWMA can be easily updated in the following way.
$$\hat{y}_t = \theta \hat{y}_{t-1} + (1 - \theta) \bar{y}_{t-1}.$$  

A plot of the EWMA of the mean composition values is presented in Figure 8.19.

Figure 8.19: EWMA of mean composition values.

Figure 8.19 is a smoothed version of Figure 8.10. With the short-term variation removed, Figure 8.19 shows even more clearly the two phenomena already described in Section 8.4.1. As was remarked there, the first peculiarity that catches the eye is the sudden jump upwards around observation 300. The second one is the slow downward movement on the left and on the right of this jump.

The explanation for the sudden jump upwards was found in the log book that is kept by the operators. In spite of the agreement not to react to bath measurements, low levels of Sn$^{2+}$ solution that were reported by bath analyses around observation 300 worried the operators. Combined with the low level of composition, it was decided to add 100 liters of new tin-plating liquid, 6.5 liters of formalin, and 6 liters of brightener. Since no immediate improvement was observed, four more liters of formalin, twenty more liters of Sn$^{2+}$ solution and 1 more liter of brightener were added the same day. Eventually, there was some reaction which is clearly visible in Figure 8.19, but also can be observed in Figures 8.9 and 8.10.
The foregoing explanation may also give a hint for explaining the slow downward trend in the composition values. Due to secondary chemical reactions, it is possible that dissolved Sn$^{2+}$ ions form Sn$^{4+}$ ions and precipitate. This may cause a slowly decreasing Sn$^{2+}$ concentration in the bath, with a similar effect on the composition observations on the products. However, for the process under consideration it is not possible to obtain accurate measurements of these bath parameters on a regular basis, so that this hypothesis cannot be verified. On the other hand, adding highly concentrated Sn$^{2+}$ solution to the bath breaks the downward trend, and may therefore be considered as an indication for a too low Sn$^{2+}$ concentration. For a complete understanding of the process mechanism, further research into this is needed.

In the database of process measurements, it was the first time that we were able to observe a long-term downward movement of the mean composition values. Previously, such movements were distorted by short-term movements, induced by making additions to the bath. Knowledge of the cause of this trend may prove to be very useful. If it is known how to influence the level of the process, knowledge of such a trend may form the basis of a control strategy. Box and Kramer (1992) have shown that control actions triggered by an EWMA crossing certain boundaries are a cost-efficient way to control IMA(1,1) processes. In the present case, this may, for example, lead to the adding of highly concentrated Sn$^{2+}$ solution to the bath, to compensate for the downward trend, thereby stabilizing the production conditions, and reducing the variation in the product measurements.

8.4.5 The special cause control chart

In the present case, the data show nonstationary behavior. However, as in situations with uncorrelated observations, special causes of variation can influence the output of the process.

Usually, in the uncorrelated case, this is narrowed down to a persistent shock in the mean level, or a persistent shock in the dispersion of the process. Suppose that the mean of the process shifts from $\mu$ to $\mu + \delta$ on time $T + 1$. The mean of deviations from the (estimated) model then shifts from 0 to $\delta$, which will lead to a higher probability of crossing one of the control limits of a control chart for the mean.

In the case of IMA(1,1) observations, an impulse shock to the level of the process by an amount of $\delta$ results in a persistent change in the level of subsequent realizations by the same amount.

As a result, a persistent change in the level of an IMA(1,1) process is
only once encountered in the residuals. Detecting a shock in the level of an IMA(1,1) process by monitoring its residuals is therefore doomed from the start, unless the shock is so large that the control limits are reached in one observation.

This does not surprise us very much, since we chose the IMA(1,1) model for its ability to capture the nonstationarity of the mean of the observations. As a consequence, the model allows the mean of the process to wander. Detecting a change in the mean by looking at the residuals is then asking for trouble.

However, a trend in the observations of an IMA(1,1) process shows up as a persistent change in the level of the residuals. The errors, made in the EWMA forecasts of mean composition values, are depicted in Figure 8.20. The control limits are computed with equal tail probabilities of $\frac{1}{2} \alpha = 0.001$.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{residuals.png}
\caption{Residuals of one step ahead forecasts.}
\end{figure}

Since we observed two persistent downward trends and one upward trend in Figure 8.10, we would, in the light of the previous arguments, expect to see some changes in the mean level of the residuals of Figure 8.20.

And indeed, there is an out-of-control signal that could be linked to a positive shift in the mean level of the residuals due to the trend around observation 300. Also, an out-of-control signal indicating a possible negative shift in the mean level of the residuals is observed. However, Figure 8.20
does not show convincingly that there is something out of the ordinary happening to the process.

The latter can be explained as follows. If we make rough estimates of the trends from the raw data, we find that for the first downward trend the level drops approximately three units in 300 observations, resulting in a decline of about 0.01 units per observation. The upward trend raises the level about six units in 60 observations, resulting in an increase of about 0.1 units per observation. The last downwards trend can also be roughly estimated as a decline of 0.01 units per observation. Hence, in Figure 8.20, we expect to see a negative shift in the level of the residuals of about 0.01 for the two downwards trends, and a positive shift in the mean of 0.1 due to the upward trend. Since the sample standard deviation of the residuals equals $\hat{\sigma}_e = 1.0479$, the change in the level of the residuals induced by the downwards trends is approximately $-0.01\sigma_e$, and the change in the level induced by the upward trend equals approximately $0.1\sigma_e$. Since these shifts are small compared to the standard error of the residuals, a regular Shewhart control chart is not the proper tool to detect such a shift.

The ARL equals 499.74 if $\delta = -0.01$ and $\alpha = 0.002$, so that, on average, 500 observations are needed before a shift of the size of the slow downward shifts is detected for the first time with a Shewhart chart such as Figure 8.20. The value of ARL(0.1) equals 475.15 if $\alpha = 0.002$ so that detecting a shift of the size of the upward shift takes, on average, about 475 observations. The ARL curve of a Shewhart chart with $\alpha = 0.002$ is depicted in Figure 8.21.

The computations above illustrate the well-known disadvantage of the Shewhart control chart that it is not very efficient in detecting small shifts in the mean. With additional run rules or warning limits its performance for detecting small shifts can be improved (see Does and Schriever (1992)). However, there are alternatives such as the EWMA and the CUSUM control chart. These control charts are more efficient in detecting small shifts in the mean.

The most common way to judge a CUSUM chart is using a Decision Interval Scheme (DIS) (see Section 5.1). Both the CUSUM for detecting positive shifts in the mean

$$S_{H_i} = \max [0, S_{H_{i-1}} + z_i - k]$$

(where $z_i$ is the standardized observation at time $i$, and $k$ is the reference value) and the CUSUM for detecting negative shifts
are compared to some value of the decision interval \( h \). One of the sums exceeding \( h \) is an indication for an out-of-control situation. The values of \( k \) and \( h \) can be chosen such that the two sided CUSUM is as efficient as possible (in terms of the ARL) in detecting a predescribed shift in the mean, while maintaining a certain value of the in-control ARL. In Figure 8.22 it is illustrated how \( k \) can be chosen to design the CUSUM for efficiently detecting a shift of size \( 0.1\sigma_\varepsilon \), while maintaining an in-control ARL of 500.

For each \( k \in [0, 1] \) we computed what value of \( h \) is needed to maintain an in-control ARL of 500. Subsequently, for this choice of \( h \) and \( k \) the ARL for \( \delta = 0.1\sigma_\varepsilon \) was computed. In Figure 8.22 these ARL(0.1) values are depicted against \( k \).

The out-of-control ARL is minimized for \( k = 0.055 \). The corresponding value of \( h \) equals 19.025, and ARL(0.1) = 237.7. Hence, this CUSUM chart will, on average, detect a shift of \( 0.1\sigma_\varepsilon \) twice as fast as a Shewhart chart. Note that this value of \( k \) agrees reasonably well with the general recommendation to set \( k \) equal to half the size of the shift we want to detect. The optimal value of \( k \) being slightly larger than half the size of the shift the chart is designed for is in accordance with the results obtained by
8.4. A NEW CONTROL STRATEGY

Figure 8.22: ARL(0.1) as a function of $k$, with $h$ such that ARL(0) = 500.

Gan (1991), who noted that this phenomenon occurs for shifts smaller than 0.2$\sigma$. In Figure 8.23, the ARL curve of a CUSUM chart with $k = 0.055$ and $h = 19.025$ is depicted.

The ARL values of the CUSUM are evaluated using the Fredholm-integral approach that is discussed in Chapter 5 and in Appendix B.

Comparing Figure 8.23 with Figure 8.21 shows that the ARL of the CUSUM is also smaller than the ARL of the Shewhart chart for other small $\delta$. However, numerical calculations show that the ARL curves of the two control charts intersect at $\delta = 1.7\sigma$. The corresponding ARL value is 12.26. For shifts in the mean larger than 1.7$\sigma$ the ARL of the Shewhart chart is smaller than the ARL of the CUSUM chart. The value of $\text{ARL}_{\text{CUSUM}} - \text{ARL}_{\text{Shewhart}}$ is maximized for $\delta = 2.8\sigma$. The ARL of the Shewhart chart then equals 2.6, whereas the ARL of the CUSUM equals 7.5. For larger values of $\delta$ this difference reduces to zero since both ARL curves converge to one.

Hence, designing the CUSUM chart in such a way that it is as sensitive as possible for detecting small shifts of size 0.1$\sigma$ in the mean results in smaller sensitivity for larger shifts as compared to the Shewhart chart.
The CUSUM chart for the mean composition values is depicted in Figure 8.24.

An out-of-control signal on the high side is observed at observation 331, indicating the upward trend mentioned earlier. This control chart shows more convincingly than the Shewhart chart that the trend in the data distorts the behavior of the one step ahead forecasts.

The previous analysis raises the question of how the model identification process was influenced by the presence of a deterministic trend. To investigate this, the data were detrended and the sample autocorrelation function of the resulting series was studied. The nonstationary behavior still appeared to be present. The estimate of the MA parameter changed from 0.811 to 0.878.

In order to estimate \( \theta \) and the two trend parameters simultaneously, two dummy variables were added to model (8.5), one for the slow (linear) downward trends, and one for the steeper upward (linear) trend. The dummies both start from zero, and are increased by one for each observation where the trend is supposed to be active.

This resulted in the following fitted model
8.4. A NEW CONTROL STRATEGY

8.4.6 Implementation and results

During the four weeks of the experiment, the mean thickness and composition observations were closely monitored. In the previous subsection we discussed in detail how this was done for the composition data. The thickness data were monitored roughly in the same way, after making a log transform in order to be able to assume normality. The variability in the measurements was monitored using regular $S$-charts, since there appeared to be no correlation in subsequent sample standard deviations.

Studying the data as described in this chapter provided valuable information that helped to achieve a better understanding of the process. In practice, however, the procedure is too laborious to be implemented,
and requires knowledge of statistical tools that did not come up in the education of the operators. A Shewhart chart of (means of) observations is much easier to work with than a CUSUM chart of residuals of one-step ahead predictions. Therefore, the following procedure is now used at Philips Stadskanaal.

From the experiment it became clear that the output of the process will slowly deteriorate if the bath is not replenished regularly. Not replenishing for a period of two weeks will bring output measurements outside specification limits. Since the bath is filtered once a week, this is a natural moment to take a sample and have it analyzed accurately by the laboratory department. Based on these measurements, the bath is replenished in such a way that all concentrations fall amply within their limits. For the rest of the week, apart from daily additions of formalin and brightener, additions are only allowed if this follows from the Out-of-Control Action Plan (OCAP), which is integrated in the automated SPC software. In the OCAP the process knowledge of the operators and chemical engineers as well as process supervisors is combined into a set of questions which initiate a systematic search for the cause of the out-of-control observations, and advise a remedy that in most cases can be applied by the operator to remove the cause that is responsible for the out-of-control signal.

The OCAP is triggered by out-of-control signals on regular Shewhart charts with widened control limits to allow for a slight wandering of the mean. This type of process monitoring is applicable since only relatively short series (one week’s data) are considered so that in ‘normal’ situations only a small deviation due to common causes in the mean will occur. We acknowledge that this approach could be improved upon, given the serial correlation present. However, in practice a balance must be struck between what is possible and what is optimal.

The result of the new replenishing strategy is illustrated by Figures 8.25 and 8.26. Figure 8.25 shows layer thickness measurements of one weeks production before changing the replenishment strategy. The production data of the same week one year later, after changing the replenishment strategy, is depicted in Figure 8.26.

From these Figures, it can be observed that the variation in the individual observations decreased considerably, and that there is a substantial decline in the number of outliers. The outcomes of the process are much more predictable. In addition, before the experiment, the mean level of layer thickness was raised to ensure that no individual measurement was found below 1μm. In the new situation, it was possible to reduce the mean level of layer thickness, without observing layer thicknesses below
8.4. A NEW CONTROL STRATEGY

![Figure 8.25: Layer thickness measurements before experiment.](image)

![Figure 8.26: Layer thickness measurements after experiment.](image)
1\mu m. This can be illustrated by comparing the average of the observations of Figure 8.25 (15.2350\mu m) to the average of Figure 8.26, which equals 9.3990\mu m.

Recall from Section 8.1 that the assignment of the PAT was to improve the solderability of the diodes. Solderability is to a large extent determined by the quality of the tin/lead layer that is applied to the diodes. Therefore, the efforts of the PAT were aimed at improving the quality of the tin/lead layer, by controlling the quality characteristics layer thickness and layer composition. The PAT succeeded in reducing the variation in these quality characteristics of the tin/lead layer, while reducing the cost of running the process of tin-plating. As a result of their achievements, the number of customer complaints regarding bad solderability of the diodes decreased to zero.