Enabling Knowledge Discovery in Multi-Objective Optimizations of Worker Well-Being and Productivity

Aitor Iriondo Pascual 1,*, Henrik Smedberg 1, Dan Högberg 1, Anna Syberfeldt 1 and Dan Lämkull 2

1 Virtual Engineering Research Environment, School of Engineering Science, University of Skövde, 541 28 Skövde, Sweden; henrik.smedberg@his.se (H.S.); dan.hogberg@his.se (D.H.); anna.syberfeld@his.se (A.S.)
2 Advanced Manufacturing Engineering, Volvo Car Corporation, 405 31 Göteborg, Sweden; dan.lamkull@volvocars.com
* Correspondence: aitor.iriondo.pascual@his.se

Abstract: Usually, optimizing productivity and optimizing worker well-being are separate tasks performed by engineers with different roles and goals using different tools. This results in a silo effect which can lead to a slow development process and suboptimal solutions, with one of the objectives, either productivity or worker well-being, being given precedence. Moreover, studies often focus on finding the best solutions for a particular use case, and once solutions have been identified and one has been implemented, the engineers move on to analyzing the next use case. However, the knowledge obtained from previous use cases could be used to find rules of thumb for similar use cases without needing to perform new optimizations. In this study, we employed the use of data mining methods to obtain knowledge from a real-world optimization dataset of multi-objective optimizations of worker well-being and productivity with the aim to identify actionable insights for the current and future optimization cases. Using different analysis and data mining methods on the database revealed rules, as well as the relative importance of the design variables of a workstation. The generated rules have been used to identify measures to improve the welding gun workstation design.

Keywords: ergonomics; digital human modeling; productivity; simulation; optimization; knowledge discovery

1. Introduction

Simulation is widely used in industries such as the automotive industry because it can efficiently create, test, and optimize the design of products and production systems in the virtual world without any need to create, test, and optimize prototypes in the physical world [1–3]. Using simulation saves time and money and allows a more thorough investigation of the solution space. These improvements in productivity have led to simulation being used in the design of workstations [4,5]. Simulation tools using digital human modeling (DHM) are also used in designing workstations to assess workers’ well-being [6].

However, simulations for optimizing productivity and simulations for assessing worker well-being are usually performed by people with different roles (production engineers solve productivity problems, while ergonomics specialists consider worker well-being). They have different focuses or objectives and use different tools. This can create silo effects, leading to slow development processes and suboptimal solutions.

Productivity and worker well-being often go hand in hand, because improving working conditions usually increases productivity [7–10]. However, sometimes the goals of productivity and well-being are at odds. Companies need to find and implement solutions in their production facilities to maintain profitability, output, and quality, as well as the well-being of workers. Previous research has identified the core elements of DHM tools and suggested a structured process for applying DHM tools in the design and development
process [11–13] at the design level of a workstation [14]. There are also applied optimization techniques to find design solutions that improve well-being and productivity [15–17]. But while there are tools available to optimize worker well-being or productivity, there are no tools available to consider both productivity and well-being while also considering the anthropometric diversity of workers.

In addition, studies often focus on finding the best solutions for a specific use case, and once solutions have been obtained and one of them has been implemented, the engineers move on to analyze the next use case. But setting up a use case and finding optimal solutions is time-consuming and requires skills and tools. Therefore, it would be beneficial to extract knowledge from existing use cases to find rules of thumb that apply to upcoming similar use cases, thereby reducing the need to perform new optimizations.

The aim of this paper is to employ the use of data mining methods to find knowledge about the solutions to a multi-objective optimization problem (MOOP) for optimizing both productivity and worker well-being, that both describes the properties of the current use case, and that potentially can be used in future use cases of similar MOOPs. Knowledge discovery has successfully been used in real-world cases for decision making in the literature, recently in [18] where the authors used clustering and association rule analysis for cantilever design problems with three and four objectives. In [19] the authors used machine learning to find relationships between the variables and objectives in a multi-objective problem of finding sweet spots in shale-gas reservoirs.

We have used data mining to find decision rules that can extract such knowledge from optimization datasets of multi-objective optimizations of worker well-being and productivity. The database generated in a previous multi-objective optimization application study [20] was mined to find rules that can be applied to similar workstations. The rules were created with the intention of balancing worker well-being and productivity, taking into account the anthropometric diversity of workers, so as to find critical solutions that can distinguish specific populations. The analysis of the multi-objective optimization application study database is meant to demonstrate that data mining approaches and knowledge discovery, in general, can generate rules that can be applied in the design of future workstations with a consequent reduction of the effort of engineers.

2. Method

We used the database of a multi-objective optimization of worker well-being and productivity application study of a welding gun workstation [20] to apply knowledge discovery methods to obtain knowledge for future workstation designs.

2.1. Application Study

The case study represents a manual welding task within manufacturing at Volvo Cars. This task involves the use of one or more of the three welding guns available for seven welding spots. The welding gun can be grasped in different positions, and workers need to adopt different postures when welding the seven spots. Since the gun is supported by a lifting device, workers are not affected by the weight of the gun. Inertia effects from moving the guns are not considered. The welding guns are not fixed in a single position; therefore, each spot can be welded by using the guns from different sides of the workstation. In addition, the task is performed by different workers, so the welding posture can change due to the anthropometric measurements of each worker. The task is repetitive, and the workers need to perform it over a full workday. Therefore, there is a risk of work-related musculoskeletal disorders (WMSD), especially in the upper limbs. The optimization was run for both worker well-being and productivity objectives.

2.1.1. Model Definition

The virtual model was modeled in the DHM tool IPS IMMA [21] by representing the welding gun workstation by imported CAD geometries. A family of 14 manikins (7 female, 7 male) was created by using an anthropometric module [22] in the DHM tool used in
the study, based on key anthropometric variables aimed to represent the anthropometric diversity of the workers at the factory, i.e., stature and elbow height (Figure 1).

Figure 1. Virtual model of the welding gun workstation in IPS IMMA.

The welding process for each welding gun at each welding spot was simulated for the entire manikin family. Table 1 shows the values for the stature and elbow height of the manikins simulated for the use case.

Table 1. Anthropometric measures of manikins.

| Manikin Number | Stature (mm) | Elbow Height (mm) | Sex   |
|----------------|--------------|-------------------|-------|
| 1              | 1629         | 984               | Female|
| 2              | 1755         | 1091              | Male  |
| 3              | 1656         | 1020              | Female|
| 4              | 1780         | 1134              | Male  |
| 5              | 1668         | 963               | Female|
| 6              | 1794         | 1068              | Male  |
| 7              | 1800         | 1094              | Female|
| 8              | 1936         | 1221              | Male  |
| 9              | 1602         | 949               | Female|
| 10             | 1731         | 1047              | Male  |
| 11             | 1590         | 1006              | Female|
| 12             | 1717         | 1114              | Male  |
| 13             | 1457         | 875               | Female|
| 14             | 1574         | 961               | Male  |

2.1.2. Well-Being and Productivity Evaluation

We evaluated the productivity of the welding workstation by measuring the cycle time of the welding sequence from the simulation. The cycle time was calculated by considering the actions of the workers performing the sequence. These actions included both value-adding operations (e.g., the time to weld a welding spot) and non-value-adding operations (e.g., changing the welding gun, changing the welding side, and moving between welding spots).

We used the Rapid Upper Limb Assessment (RULA) method [23] to evaluate the risk of WMSDs. We considered RULA appropriate since the work stresses mainly involve
postural stresses on the upper limbs, as the weight of the welding guns is supported by a lifting device. For each posture assessed, RULA gives a risk score from 1 to 7, which results in four action levels. A score of 1 to 2 is acceptable, a score of 3 to 4 suggests changes may be required, 5 to 6 indicates changes will soon be required, and 7 indicates that changes are required immediately.

2.1.3. Mathematical Modeling of Optimization

The optimization model considers both worker well-being and productivity objectives, that is, it is a multi-objective optimization model. The indices, parameters, variables, and objectives of the optimization model are shown in Table 2.

Table 2. Indices, parameters, variables, and objectives of the optimization model.

| Indices | Parameters |
|---------|------------|
| \(w = 1 \ldots W\) | Welding spots | \(TW\) | Welding time (s) |
| \(g = 1 \ldots G\) | Welding guns | \(TG\) | Time to change welding gun (s) |
| \(s = 1 \ldots S\) | Welding sides | \(TS\) | Time to change welding side (s) |
| \(m = 1 \ldots M\) | Manikins | \(TF\) | Time to move to a far position (s) |
| \(sq = 1 \ldots SQ\) | Welding sequence | \(TN\) | Time to move to a near position (s) |

| Variables | | |
|----------|----------|
| \(X_w\) | Welding spot sequence | \(PG_{gr}\) | Previous gun: 1 if different, 0 if same |
| \(Y_g\) | Welding gun used at each welding spot | \(PS_{sq}\) | Previous side: 1 if different, 0 if same |
| \(Z_m\) | Welding side at each welding spot | \(PF_{gsw}\) | 1 if previous spot is far, 0 if near |
| \(Z_sq\) | Welding side at each welding spot | \(PN_{sq}\) | 1 if previous spot is near, 0 if far |

| Objectives | | |
|------------|----------|
| \(CT\) | Cycle time of welding process (s) | \(ER_m\) | RULA score for a manikin on a side with a welding gun at a welding spot |
| \(ER_w\) | Average RULA score per manikin in the welding process | \(ER\) | Average RULA score of all manikins in the welding process |

A multi-objective optimization of the average of the RULA scores for all the manikins and the cycle time was performed. The mean RULA score of the 14 manikins was calculated as a mean value for all the manikins using all welding spots. The risk of WMSDs is therefore calculated by a single objective:

\[
MIN \overline{ER} = \frac{\sum_{m=1}^{M} \sum_{sq=1}^{SQ} ER_{msgw}}{M}
\]  

(1)

The cycle time was calculated as the sum of the welding time at each welding spot and the time to change welding gun, welding spot, and welding side:

\[
MIN \; CT = \sum_{sq=1}^{SQ} TW + \sum_{sq=2}^{SQ} PG_{sq} \cdot TG + \sum_{sq=2}^{SQ} PS_{sq} \cdot TS + \sum_{sq=2}^{SQ} PF_{gsw} \cdot TF + \sum_{sq=2}^{SQ} PN_{sq} \cdot TN
\]  

(2)

2.1.4. Optimization Method

Many real-world optimization problems involve multiple conflicting objectives and therefore lead to multiple Pareto-optimal solutions, that is, the optimal solution for one optimization objective may not be optimal for another objective [24]. Therefore, solving a MOOP is about balancing the trade-offs among conflicting objectives. This complexity also affects the process of obtaining optimal solutions while performing optimization. Due to their population-based nature, multi-objective evolutionary algorithms are often used for solving MOOPs. The evolutionary algorithm NSGA-II was used to optimize
this application study because of its efficiency in multi-objective optimizations [25]. The parameters used for the optimization are presented in Table 3.

### Table 3. Optimization algorithm configuration.

| Optimization Algorithm | NSGA-II          |
|------------------------|------------------|
| Population size        | 150              |
| Child population size  | 150              |
| Tournament size        | 2                |
| Mutation operator      | Polynomial       |
| Mutation probability   | 0.2              |
| Crossover probability  | 0.9              |
| Crossover operator     | SBX              |
| Maximum iterations     | 25,000           |

2.2. Knowledge Discovery

We sought to extract general knowledge from the solutions to the MOOP. A MOOP solution can be seen as inhabiting two different spaces: the decision space, where the decision variables exist, and the objective space where the objective values exist. While optimization algorithms drive the search to find higher performing solutions in the objective space, data mining methods can be used to find knowledge about preferred solutions in terms of the decision space. To generate knowledge, we defined the preference by analyzing the solutions, and then applied a data mining method to generate decision rules to describe those preferences. The process of knowledge discovery is done in several steps: (1) data filtering, (2) data clustering, (3) data visualization, (4) rule extraction, and (5) knowledge interpretation.

The first step, data filtering, was done by removing the duplicates in the database to remove any bias in the resulting knowledge (Figure 2). In the second step, data clustering, we identified four principal objectives: lowest cycle time, lowest average RULA score for all manikins, balance between average RULA score for all manikins and cycle time, and worker diversity inclusion rules. In the next step, data visualization, we analyzed the database by using parallel coordinates, boxplots, and 2D and 3D scatter plots. Later, knowledge discovery was performed using influence score by rank (InfS-R) and flexible pattern mining (FPM) methods [26]. In the final step, knowledge interpretation, we analyzed the generated rules and related them to the specific implications in the application study to generate reproducible knowledge for future workstation designs.

![Flow diagram of the knowledge discovery process.](image-url)
2.2.1. FPM

FPM [27] is a recently developed method for obtaining decision rules in the decision space, based on selections made by the decision maker in the objective space. FPM is an extension of sequential pattern mining [28] and can find rules of the forms \{x_i < c_1\}, \{x_i > c_2\}, and \{x_i = c_3\} for a decision variable \(x_i\) and constant values \(c_1, c_2,\) and \(c_3\). The FPM procedure finds decision rules about the decision space that describes selections made in the objective space. To run the process, the decision maker must supply a set of selected solutions and a set of unselected solutions. The resulting FPM rules can be seen as a tuple containing the following elements: label, referring to the name of the decision variable; sign, referring to one of the signs \(<, >,\) or \(=\); value, referring to the constant value of the rule; sig., referring to the significance of the rule, that is, the percentage of the solutions in the selected set that the rule covers; and unsig., referring to the unselected significance (unsignificance), that is, the percentage of solutions in the unselected set that the rule covers. Note that since the significance and unsignificance can be found simply by counting the solutions covered by a single rule, the same can be done for interactions of two or more rules. In Section 3 we show three levels of FPM rule interactions. An interesting and descriptive FPM rule would have high significance and low unselected significance. The ratio of significance over unselected significance serves as a general metric for a good rule.

2.2.2. InfS-R

Factor screening to find which variable has the most influence in a model can be a useful tool for understanding the model and how the variables interact [29]. Recently, an approach for finding influential variables in the resulting solution set from a multi-objective optimization problem was presented in [26]. The approach defines two methods. One finds an influence score for the decision variables in terms of how they affect the solutions on the Pareto-optimal front; this method is called influence score by Pareto front (InfS-P). The other method finds an influence score for decision variables in terms of how they affect the convergence on the Pareto front by considering the rank of the solutions after non-dominated sorting, called InfS-R. Both methods divide the solutions into different selections and find FPM rules to describe the differences between them. The frequency of and significance of the different variables is then used to determine how influential they are. We use InfS-R to find the relative influence of the contribution of certain variables to generating Pareto-optimal solutions.

3. Results

The solution space from the optimization is presented in a 2D scatter plot (Figure 3) of the colliding objectives of average RULA for all manikins and cycle time.

![Figure 3. Solution space of cycle time and average RULA score ER of all manikins.](image-url)
The solutions that correspond to the non-dominated solutions represent the Pareto front for this optimization (marked by “+” in Figure 5) and are the best solutions of the optimization. These solutions are shown in Table 4.

Table 4. Results from optimization.

| Result Selected | CT (s) | E_\text{R} | Sequence |
|-----------------|--------|-------------|----------|
| Lowest CT       | 47     | 3.09        | Spot sequence: 7-1-3-2-5-4-6  
|                 |        |             | Gun sequence: 3-3-3-3-3-3-3  
|                 |        |             | Side sequence: 1-1-1-1-2-2-2  |
| Compromise between CT and E_\text{R} | 63     | 2.89        | Spot sequence: 4-5-6-7-1-2-3  
|                 |        |             | Gun sequence: 3-3-3-3-2-2-2  
|                 |        |             | Side sequence: 2-2-2-1-1-1-1  |
| Lowest E_\text{R} | 85     | 2.86        | Spot sequence: 4-5-6-7-1-2-3  
|                 |        |             | Gun sequence: 4-5-6-7-1-2-3  
|                 |        |             | Side sequence: 2-2-2-1-1-1-1  |

3.1. Data Filtering

The database of the application study was filtered by removing duplicates. Also, in order to help the data mining process, related variables were merged. In this case, the welding guns and the welding sides were converted into a single variable “gun and side” before proceeding with the data mining. To convert gun and side to the same variable, the variables were converted into a single integer per spot, GS_w. Due to the availability of the guns and sides in different welding spots, the range was different for every welding spot (Table 5).

Table 5. Corresponding value for every GS_w depending on the gun and side availability at each spot.

| w = 1 | s = 1 | s = 2 |
|-------|-------|-------|
| GS_1  | X     | GS_1  |
| GS_1  | X     | GS_2  |
| GS_1  | X     | GS_3  |
| GS_1  | X     | GS_4  |
| GS_1  | X     | GS_5  |
| GS_1  | X     | GS_6  |
| GS_1  | X     | GS_7  |

Note: X is unavailable.

3.2. Data Clustering

The selected clusters that need to be analyzed represent the (1) lowest cycle time (CT) (marked by “+”), (2) lowest average RULA score for all manikins (E_\text{R}) (marked by “.”), (3) the balance between average RULA score for all manikins and cycle time (marked by “x”), and (4) worker diversity inclusion rules. The clusters of (1) lowest CT and (2) lowest E_\text{R} were defined by selecting the solutions with the lowest scores. The cluster of (3) (balance between CT and E_\text{R}) was defined as all solutions that had a lower CT than cluster (2), and lower E_\text{R} than cluster (1) (Figure 4).
An initial analysis of the manikins was necessary to select the cluster for (4) (diversity inclusion rules) since the critical manikins were still not identified.

3.3. Data Visualization

A boxplot analysis of the average RULA score of each manikin ($ER_m$) was made to evaluate the critical manikins and consider worker anthropometric diversity (Figure 5).

Manikins 13 and 14 have higher average RULA scores than other manikins for all solutions (Figure 5). These are the manikins with the lowest stature in the population considered (Table 1). The cluster for (4) (diversity inclusion rules) was therefore formed by the best solutions for the $ER_m$ of manikins 13 and 14, $ER_{13}$ and $ER_{14}$ (marked by “x” in Figure 6).
were the selection of welding gun and side for every welding spot (\(G_S\) or \(\text{SpotW G/S}\) in figures) and the welding sequence (\(X_w\)) (Figure 7).

### 3.4. Knowledge Discovery

To obtain rules from the results of the application study, InfS-R [26] was used to determine the relative importance of the variables in the optimization. The first InfS-R ranked the non-dominated sorting for the objectives \(CT\) and \(\overline{ER}\). The variables studied were the selection of welding gun and side for every welding spot (\(G_S\) or \(\text{SpotW G/S}\) in figures) and the welding sequence (\(X_w\)) (Figure 7).

![Figure 6](image_url)  
**Figure 6.** Selection of results for cluster 4, diversity inclusion rules.

![Figure 7](image_url)  
**Figure 7.** And \(\overline{ER_{13}}\) to analyze the diversity inclusion (Figure 8).
We used the Rapid Upper Limb Assessment (RULA) method [23] to evaluate the risk of WMSDs.

The cycle time was calculated as the sum of the welding time at each welding spot and the time to change welding gun, welding spot, and welding side:

\[ \text{Cycle Time} = \sum \text{Welding Time at each spot} + \text{Time to change welding gun, welding spot, and welding side} \]

A compromise between welding spots was considered, therefore calculated by a single objective:

\[ \text{Compromise between welding spots} = \text{Minimize Cycle Time} \]

Worker diversity inclusion was analyzed by the InfS-R [17] methodology to analyze the diversity inclusion (Figure 8).

After the InfS-R analysis, an FPM analysis was run for the four clusters (1) lowest cycle time (marked by “+” in Figure 4), (2) lowest average RULA score for all manikins (marked by “−” in Figure 4), (3) balance between average RULA score for all manikins and cycle time (marked by “x” in Figure 4), and (4) worker diversity inclusion rules (marked by “x” in Figure 6). The FPM analysis was run for three levels of rule-interactions and 0.5 minimum significance and minimum support. After that, the rules were filtered to the highest ratio of significance/unsignificance to obtain the most relevant rules for the four clusters. The significance, unsignificance, and ratio of the obtained rules in the four clusters are presented in Table 6.

Table 6. Rules obtained for the four cases by FPM.

| Case                  | Filtered Rules          | Sig. (%) | Unsig. (%) | Ratio |
|-----------------------|-------------------------|----------|------------|-------|
| Lowest CT             | \( GS_1 = 4 \)          | 91.53    | 27.07      | 3.38  |
|                       | \( GS_2 > 2 \)          | 100      | 40.85      | 2.45  |
|                       | \( GS_3 = 3 \)          | 79.1     | 28.86      | 2.74  |
|                       | \( GS_1 = 4 \&\& GS_2 > 2 \&\& GS_3 = 3 \) | 75.14    | 7.58       | 9.91  |
| Compromise between CT and \( \text{ER} \) | \( GS_1 < 4 \)          | 71.29    | 68.59      | 1.03  |
|                       | \( GS_3 > 2 \)          | 64.36    | 45.51      | 1.41  |
|                       | \( X_5 > 2 \)           | 82.18    | 69.4       | 1.18  |
|                       | \( GS_1 < 4 \&\& GS_3 > 2 \&\& X_5 > 2 \) | 52.48    | 17.38      | 3.02  |
| Lowest \( \text{ER} \) | \( GS_1 = 2 \)          | 100      | 41.86      | 2.39  |
|                       | \( GS_2 = 2 \)          | 100      | 31.93      | 3.13  |
|                       | \( GS_3 = 1 \)          | 100      | 23.45      | 4.26  |
|                       | \( GS_1 = 2 \&\& GS_2 = 2 \&\& GS_3 = 1 \) | 100      | 0.73       | 136.99|
| Worker diversity inclusion | \( GS_1 = 2 \)          | 93.84    | 38.17      | 2.46  |
|                       | \( GS_2 = 2 \)          | 93.84    | 26.94      | 3.48  |
|                       | \( GS_3 = 1 \)          | 82.35    | 22.73      | 3.62  |
|                       | \( GS_1 = 2 \&\& GS_2 = 2 \&\& GS_3 = 1 \) | 75.95    | 0.32       | 237.34|
3.5. Knowledge Interpretation

The boxplot analysis showed that manikins 13 and 14 gave higher values for $\overline{ER}_m$ than the other manikins (Figure 5). It can also be seen that the minimum values for manikins 13 and 14 were higher than the maximum values for the rest of the manikins. The values for $\overline{ER}_{13}$ and $\overline{ER}_{14}$ were high due to the stature of these manikins (Table 1). This meant that the cluster for (4) (diversity inclusion rules) was defined by the solutions that have low values for manikins 13 and 14.

Once the clusters were defined, two InfS-R analyses were run to identify the relative importance of the selection of guns, sides, and welding sequence. The first analysis was run for the objectives of $CT$ and $\overline{ER}$. The variables with the highest relative importance were the selected gun and side in spots 1, 2, 3, and 7 ($GS_1$, $GS_2$, $GS_3$, and $GS_7$) and the welding sequence. The second analysis was run for the objectives of $\overline{ER}_{13}$ and $\overline{ER}_{14}$. The variables with the highest relative importance were also the selected gun and side in spots 1, 2, 3, and 7 ($GS_1$, $GS_2$, $GS_3$, and $GS_7$). However, for the objectives of $\overline{ER}_{13}$ and $\overline{ER}_{14}$, the welding sequence did not have high relative importance since the average $RULA$ values of the manikins are only affected by the gun and side used in the welding.

After the InfS-R analyses, four FPM analyses were run for the clusters (1) lowest cycle time ($CT$), (2) lowest average $RULA$ score for all manikins ($\overline{ER}$), (3) balance between average $RULA$ score for all manikins and cycle time, and (4) worker diversity inclusion rules. The rules were produced by running the FPM analyses for 3 levels of rule-interactions. The rules were filtered by selecting the ones with the highest ratio of significance/significance. The resulting rules apply in clusters (1), (3) and (4) to $GS_1$, $GS_2$, and $GS_3$, as expected from the first InfS-R study. Also, a rule was found for the welding sequence $X_5$ in the cluster (2) for the compromise between $CT$ and $\overline{ER}$. However, it has a high unsignificance and therefore is not very interesting since it does not distinguish the selection.

The cluster with the highest ratio in the created rules is (4), worker diversity inclusion. The rules define the welding gun and side used in spots 1, 2, and 3, as expected from the second InfS-R analysis (Figure 8). The rules $GS_1 = 2$, $GS_2 = 2$ and $GS_3 = 1$ have a ratio of 2.46, 3.48, and 3.62 and a significance of 93.84%, 93.84%, and 82.35%, respectively, and a combined ratio of 237.34 and significance of 75.85%. This means that these rules strongly define cluster (4) against the rest of the solutions, with an unsignificance of 0.32%. If the decision maker was to select a solution of (4) (worker diversity inclusion), then welding gun 2 should be used in spots 1 and 2, and welding gun 1 should be used in spot 3. This would mean changing the welding gun at least twice in the welding sequence, with a consequent increase in $CT$.

The same rules were applied in cluster (2) to find the lowest average $RULA$ score for all manikins ($\overline{ER}$). While in cluster (2) the ratio for the rules combined was lower than for cluster (4), in this case the significance of each of the three rules was 100%. This means that all the solutions that have the lowest average $RULA$ score for all manikins require using welding gun 2 in spots 1 and 2 and welding gun 1 in spot 3. This is also reflected in the lowest $\overline{ER}$ solution of the Pareto front (Table 4).

For the cluster (1) (lowest cycle time—$CT$), the rules generated are $GS_1 = 4$, $GS_2 > 2$, and $GS_3 = 3$, which have a significance of 91.53%, 100%, and 79.1% and a ratio of 3.38, 2.45, and 2.74, respectively. With a significance of 100%, rule $GS_2 > 2$ indicates that welding gun 3 (on any side) must be used to obtain low cycle time scores at welding spot 2. This collides with the rules obtained for clusters (2) and (4), which show that for a low average $RULA$ score for all manikins ($\overline{ER}$), it is necessary to use welding gun 1. Also, rules $GS_1 = 4$ and $GS_3 = 3$ for a low cycle time indicates that welding gun 3 should also be used in spots 1 and 3. The lowest $CT$ solution in the Pareto front (Table 4) also shows that using welding gun 3 in all welding spots gives the lowest cycle time. This is due to the time saved by not changing the welding gun during the process.

Of the four clusters, the rules for cluster (3) have the lowest ratio. These rules have a high unsignificance due to the difficulty in finding rules that describe this cluster. Rule
GS1 < 4 only discards using welding gun 3 on side 2 for spot 1, and rule GS3 > 2 only discards using welding gun 1 in spot 3. In the case of rule X5 > 2, the rule only defines that spot 5 should be welded after two other spots. These three rules have a ratio of 1.03, 1.41, and 1.18 respectively, and a ratio of 3.02 when combined. With these low ratios, the rules do not describe the cluster (3) and distinguish it from the other solutions. This could be because defining the balance between CT and ER cannot be done by straightforward descriptions but requires a mix of the rules of clusters (1) and (2).

4. Discussion

In this study we aimed to use data mining in multi-objective optimizations of worker well-being and productivity to discover knowledge that could be applied in future workstation designs. The initial optimization showed that the consideration of cycle time together with RULA scores allowed analysis of the impact of different configurations of the welding sequence. Also, optimization allowed consideration of the anthropometric diversity of the workers, helping workstation designers accommodate the diversity of the workforce. The use of different analysis and data mining methods on the database generated from the multi-objective optimization of worker well-being and productivity allowed rules to be discovered, and showed the relative importance of the design variables of the workstation.

The boxplot analysis of the individual manikins average RULA scores (ERm) showed that manikins 13 and 14, the manikins representing the lowest stature percentiles of female and male populations, had the highest RULA scores (ER13 and ER14). The workstation design is clearly more adapted to high stature percentile populations than to low stature percentile populations. Redesigning the workstation to be at a lower height could benefit shorter workers, improving their posture while welding. Due to this result for manikins 13 and 14, we defined a cluster for the best solutions for manikins 13 and 14.

We performed two Infs-R analyses, one for the CT and ER objectives, and a second one for ER13 and ER14 to consider the critical manikins. In both cases, GS1, GS2, GS3 and GS7 had the highest relative importance. We realized that the lack of rules that included GS1, GS5, and GS6 implied that those spots did not allow the use of welding guns 1 and 2 (Table 5). Therefore, it was impossible to improve any objective by changing only the side where welding gun 3 was used, leading to a low diversity of solutions. Removing some of the constraints that impede using welding guns 1 and 2 in spots 4, 5, and 6 would increase the diversity in the solutions, allowing new solutions to be found that could benefit both worker well-being and productivity.

After using Infs-R, FPM was used in the four clusters to find rules. The rules found apply to the actual design of the workstation, that is, the workstation that is constrained in spots 4, 5 and 6. The rules, therefore, apply mostly in the selection of gun and side at spots 1, 2, and 3. There are no rules that include the selection of gun and side for welding spot 7 since welding gun 3 obtained both better CT and ER results at that spot. The rules for clusters (3) and (4) contradict the rules of cluster (1). This means that when it comes to the actual design of the workstation, there is a clash between productive solutions (solutions with low cycle time) and solutions relating to worker well-being and inclusivity (average RULA score of the manikins). This clash does not imply that one solution should be chosen without regard for the other (equally valid) objective. Instead, the design of the workstation should be modified. In this case, welding spots 1, 2, and 3 appeared in the rules of all clusters creating the conflict between objectives. Therefore, modifying the design so that other design solutions are generated could benefit both objectives at the same time.

The knowledge discovered in this study could be applied in future designs of workstations, where workstation designers should try to not constrain welding positions for different guns. In order to keep the cycle time to a minimum, the welding gun should be changed as few times as possible. To keep low RULA average scores, the workers should be able to use the welding gun in different positions depending on their anthropometric measures. Therefore, it is critical to design the workstation for specific welding guns taking into account the available welding positions. This would allow optimizations with higher
was done by H.S. and A.I.P. H.S. also performed the data curation. The supervision and the review and editing were performed by D.H., A.S. and D.L. D.H. and A.S. also contributed with the funding acquisition. D.L. also provided the resources for the use case. All authors have read and agreed to the published version of the manuscript.

Funding: This work has received support from ITEA3/Vinnova in the project MOSIM (2018-02227), and from Stiftelsen för Kunskaps- och Kompetensutveckling within the Synergy Virtual Ergonomics (SVE) project (2018-0167) and the Virtual Factories—Knowledge-Driven Optimization (VF-KDO) research profile (2018-0011). This support is gratefully acknowledged.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors would like to thank the organizations participating in the projects associated to this work. Their support is gratefully acknowledged.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

5. Conclusions

Using optimization algorithms to find optimized workstation designs allows the solution space to be explored by a strategic search through feasible solutions without manually processing each of all possible configurations. The use of the Infs-R and FPM methods in the database of all solutions generated in the optimization provides a deeper understanding of the behavior of the workstation design, allowing engineers to identify critical factors in the workstations and later improve them. By using these methods in the welding gun use case, we were able to identify the critical design improvements necessary to improve both workers’ well-being and productivity. We discovered that the workstation design constrained welding guns 1 and 2, removing the constraints could provide better solutions for welding spots 4, 5, and 6. This could lead to better solutions for CT and $ER$, and also considers anthropometric diversity in the workstation. The knowledge discovered in this study could be applied in the design of future workstations, so that engineers can avoid constrained positions of welding guns and generate better workstation design solutions to improve productivity and worker well-being in factories.

The use of knowledge discovery for multi-objective optimizations of worker well-being and productivity can be used in relation to different workstation designs. Such knowledge can help to engineers find good design solutions for all future workstations, including other types of workstations such as assembly-line workstations.

Knowledge discovery requires engineers to have some expertise in performing multi-objective optimizations and extracting the knowledge from the databases of the optimizations. In order to further support engineers, the optimization setup and knowledge discovery process should be implemented in a digital tool that considers users’ expertise and guides the users through the entire process.
References

1. Kuhn, W. Digital Factory—Simulation Enhancing the Product and Production Engineering Process. In Proceedings of the 2006 Winter Simulation Conference, Monterey, CA, USA, 3–6 December 2006; pp. 1899–1906. [CrossRef]

2. Fisher, D.L.; Lee, J.D.; Rizzo, M.; Caird, J. Handbook of Driving Simulation for Engineering, Medicine, and Psychology; CRC Press: Boca Raton, FL, USA, 2011.

3. Oppelt, M.; Urbas, L. Integrated virtual commissioning an essential activity in the automation engineering process: From virtual commissioning to simulation supported engineering. In Proceedings of the IEEE Industrial Electronics Society, Dallas, TX, USA, 29 October–1 November 2014; pp. 2564–2570. [CrossRef]

4. Wang, Q.; Chatwin, C. Key issues and developments in modelling and simulation-based methodologies for manufacturing systems analysis, design and performance evaluation. Int. J. Adv. Manuf. Technol. 2004, 25, 1254–1265. [CrossRef]

5. Ore, F.; Vemula, B.R.; Hanson, L.; Wiktorsson, M. Human—Industrial Robot Collaboration: Application of Simulation Software for Workstation Optimisation. Procedia CIRP 2016, 44, 181–186. [CrossRef]

6. Scataglini, S.; Paul, G. Key issues and developments in modelling and simulation-based methodologies for manufacturing systems analysis, design and performance evaluation. Int. J. Adv. Manuf. Technol. 2004, 25, 1254–1265. [CrossRef]

7. Falck, A.-C.; Örtengren, R.; Högborg, D. The impact of poor assembly ergonomics on product quality: A cost-benefit analysis in car manufacturing. Hum. Factors Ergon. Manuf. Serv. Ind. 2010, 20, 24–41. [CrossRef]

8. Widiana, I.; Sumetri, N.W.; Sutapa, I.K. Ergonomic Work Station Design to Improve Workload Quality and Productivity of the Craftsmen. J. Phys. Conf. Ser. 2018, 953, 12091. [CrossRef]

9. Green, R.F. A Generic Process for Human Model Analysis; SAE Technical Paper 2000-01–2167; SAE International: Warrendale, PA, USA, 2000. [CrossRef]

10. Blomé, M.W.; Dukic, T.; Högborg, D. Guide and documentation system to support digital human modeling applications. Int. J. Ind. Ergon. 2006, 36, 17–24. [CrossRef]

11. Reed, M.P.; Faraway, J.; Chaffin, D.B.; Martin, B.J. The HUMOSIM Ergonomics Framework: A New Approach to Digital Human Simulation for Ergonomic Analysis; SAE Technical Paper 2006-01–2365; SAE International: Warrendale, PA, USA, 2006. [CrossRef]

12. Battini, D.; Faccio, M.; Personna, A.; Sgarbossa, F. New methodological framework to improve productivity and ergonomics in assembly system design. Int. J. Ind. Ergon. 2011, 41, 30–42. [CrossRef]

13. Mura, M.D.; Dini, G. Optimizing ergonomics in assembly lines: A multi objective genetic algorithm. CIRP J. Manuf. Sci. Technol. 2017, 29, 31–45. [CrossRef]

14. Högberg, D.; Lämkull, D.; Luque, E.P.; Syberfeldt, A.; Hanson, L. Optimization of productivity and worker well-being by using a multi-objective optimization framework. IEEE Trans. Hum.–Mach. Syst. 2019, 49, 440–448. [CrossRef]

15. Sana, S.S.; Mateus, H.O.; Arrieta, F.A.G.; Acevedo-Chedid, J. Application of genetic algorithm to job scheduling under ergonomic constraints in manufacturing industry. J. Ambient Intell. Humaniz. Comput. 2018, 10, 2063–2090. [CrossRef]

16. Battini, D.; Faccio, M.; Personna, A.; Sgarbossa, F. New methodological framework to improve productivity and ergonomics in assembly system design. Int. J. Ind. Ergon. 2011, 41, 30–42. [CrossRef]

17. Dini, G.; Pardalos, P., Eds.; Springer: Boston, MA, USA, 2000. [CrossRef]

18. Ahmad, J.; Bechar, A.; Riemer, R. Simulation-Based Optimization Methodology for a Manual Material Handling Task Design That Maximizes Productivity While Considering Ergonomic Constraints. IEEE Trans. Hum.–Mach. Syst. 2019, 49, 440–448. [CrossRef]

19. Sato, Y.; Izu, K.; Yamada, T.; Nishiwaki, S. Data mining based on clustering and association rule analysis for knowledge discovery in multiprojective topology optimization. Expert Syst. Appl. 2019, 119, 247–261. [CrossRef]

20. Tahmasebi, P.; Javadpour, F.; Sahimi, M. Data mining and machine learning for identifying sweet spots in shale reservoirs. Expert Syst. Appl. 2017, 88, 433–447. [CrossRef]

21. Pascual, A.; Högborg, D.; Lämkull, D.; Luque, E.P.; Syberfeldt, A.; Hanson, L. Optimization of productivity and worker well-being by using a multi-objective optimization framework. IEEE Trans. Occup. Ergon. Hum. Factors 2021, 9, 143–153. [CrossRef]

22. Högberg, D.; Hanson, L.; Bohlin, R.; Carlson, J.S. Creating and shaping the DHM tool IMMA for ergonomic product and production design. Int. J. Digit. Hum. 2016, 1, 132–152. [CrossRef]

23. McAtamney, L.; Corlett, E.N. RULA: A survey method for the investigation of work-related upper limb disorders. Appl. Ergon. 1993, 24, 91–99. [CrossRef]

24. Sayin, S. Multi-objective optimization and decision support systems. In Encyclopedia of Optimization, 2nd ed.; Floudas, C.A., Pardalos, P., Eds.; Springer: Boston, MA, USA, 2009.

25. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. 2002, 6, 182–197. [CrossRef]

26. Smedberg, H.; Bandaru, S. Finding influential variables in multi-objective optimization problems. In Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Cancun, ACC, Australia, 1–4 December 2020; pp. 173–180. [CrossRef]

27. Bandaru, S.; Ng, A.H.; Deb, K. Data mining methods for knowledge discovery in multi-objective optimization: Part B—New developments and applications. Expert Syst. Appl. 2017, 70, 119–138. [CrossRef]

28. Agrawal, R.; Srikan, R. Mining sequential patterns. In Proceedings of the Eleventh International Conference on Data Engineering, Taipei, Taiwan, 6–10 March 1995; pp. 3–14. [CrossRef]

29. Kleijn, J.P.C. Factor screening in simulation experiments: Review of sequential bifurcation. In Advancing the Frontiers of Simulation: A Festschrift in Honor of George Samuel Fishman; Alexopoulos, C., Goldsman, D., Wilson, J.R., Eds.; Springer: Boston, MA, USA, 2009; pp. 153–167. [CrossRef]