MDPI

Article

Software Sensor for Activity-Time Monitoring and Fault Detection in Production Lines

Tamas Ruppert and Janos Abonyi *

MTA-PE "Lendület" Complex Systems Monitoring Research Group, Department of Process Engineering, University of Pannonia, Egyetem u. 10, POB 158, H-8200 Veszprém, Hungary; ruppertt@fmt.uni-pannon.hu

* Correspondence: janos@abonyilab.com

check for updates

Abstract: Industry 4.0-based human-in-the-loop cyber-physical production systems are transforming the industrial workforce to accommodate the ever-increasing variability of production. Real-time operator support and performance monitoring require accurate information on the activities of operators. The problem with tracing hundreds of activity times is critical due to the enormous variability and complexity of products. To handle this problem a software-sensor-based activity-time and performance measurement system is proposed. To ensure a real-time connection between operator performance and varying product complexity, fixture sensors and an indoor positioning system (IPS) were designed and this multi sensor data merged with product-relevant information. The proposed model-based performance monitoring system tracks the recursively estimated parameters of the activity-time estimation model. As the estimation problem can be ill-conditioned and poor raw sensor data can result in unrealistic parameter estimates, constraints were introduced into the parameter-estimation algorithm to increase the robustness of the software sensor. The applicability of the proposed methodology is demonstrated on a well-documented benchmark problem of a wire harness manufacturing process. The fully reproducible and realistic simulation study confirms that the indoor positioning system-based integration of primary sensor signals and product-relevant information can be efficiently utilized in terms of the constrained recursive estimation of the operator activity.

Keywords: recursive estimation; performance monitoring; indoor positioning system; paced conveyor; early warning systems

1. Introduction

In the age of digital transformation, human operators are still applied in manufacturing processes. The Operator 4.0 concept aims to create human-cyber-physical production systems (H-CPPS) that improve the abilities of the operators' thanks to the dynamic interaction between humans and production systems [1]. Smart sensors are key components of CPPS solutions [2]. Model-based production control and performance monitoring require accurate information concerning the activity times of the operators. Handling human factors is a challenging problem in terms of both cellular manufacturing [3] and human-robot interaction [4]. Usually operator activity is monitored by computer vision-based motion detection systems and Radio Frequency IDentification (RFID)-based object tracking [5]. Context-aware systems require unobtrusive sensors to track each step of the performed task and present the worker with the information needed at any given moment [6]. As wearable sensors are becoming more common, their utilization is also becoming more attractive [7]. However, hand motion-based activity recognition is still challenging [8] and requires the application of advanced machine learning algorithms [9]. As this brief overview shows as well, the tracking of operator activity

is a difficult, highly infrastructure-demanding task which should utilize information stream fusion approaches to improve the robustness of the algorithms [10].

Tracing hundreds of primary activities is critical due to the enormous variability and complexity of products. As every operator performs sequentially a specific set of actions over a period of time, our goal is to develop a sensor system that continuously estimates the time consumption of these elementary activities. We model the time consumptions of these actions by activity time models and compare the estimated activity times to the performance of operators and generate early warnings when their productivity decreases.

For the cost-effective and robust measurement of assembly times, sensors were developed to record the timestamps related to the activity when the components are pushed into the fixtures by operators. As the activities of operators depend on the type and number of the built-in components, the production flow is tracked by an indoor positioning system (IPS). For the localization of the products and identification of the status of the conveyor system, Ultra-Wide band (UWB) IPS technology is applied with its low energy demand for transmitting information over a broad bandwidth (>500 MHz) and its accuracy with the range of 30–50 cm, which is significantly better than the one-meter uncertainty of Bluetooth Low Energy (BLE)-based solutions [11,12].

To integrate measurements originating from the IPS, a varying number (10–100) of active or passive fixture sensors, and other information sources of the production management system, a multi-sensor data fusion (MSDF) algorithm has been developed. Multiple sensors provide redundancy enabling the robust recursive estimation of the unmeasured primary activity times of the operators. To constrain the model parameters to lie within a reliable region and incorporate important *a priori* knowledge concerning the activity times, the estimated parameters were optimally projected on to a set of linear constraints by quadratic programming [13]. This central estimation enhances the confidence of the nominal model which improves the performance of fault detection based on the reconciliation of the local measurements.

The development of the proposed fault-detection algorithm is motivated by the analysis of an industrial wire harness manufacturing process which is a typical complex modular product manufacturing system [14,15]. To ensure our results are fully reproducible, only openly available information on wire harness manufacturing technologies was utilized during the development of the realistic case study. To stimulate further research, the resultant algorithm of the developed model of the manufacturing system and the details of the products and sensor placements are publicly available on the website of the authors (https://www.abonyilab.com/soft-sensors).

The remaining part of the paper is structured as follows. The developed IIoT-based sensor system is shown in Section 2. The applicability of the proposed activity-time estimation algorithm is demonstrated in Section 3. Based on the findings and discussions reported there, conclusions are drawn in Section 4.

2. Software Sensor for Activity-Time Monitoring

In the present section, first the conveyor and the modular production systems are characterized, then the fixture sensors and the indoor positioning system as information sources are described. This is followed by the mathematical formulation of the multi sensor data fusion-based recursive estimation model and finally by the local estimation and monitoring with regard to the activity times of operators.

2.1. Problem Definition—Evaluation of Activity Times on the Paced Conveyor

The development of the proposed fault-detection algorithm is motivated by the analysis of an industrial wire harness manufacturing process. Wire harnesses are produced by a typical complex modular production system [14,15]. The crucial part of the studied wire harness manufacturing system is a similar conveyor system as shown in Figure 1. The motion of the conveyor is paced and cyclic in nature. At the beginning of the cycles, every station proceeds to the next position. The operators might work ahead of schedule or be delayed. According to the open-station concept, when the operator

does not finish his or her job, he or she can move with the product to the next station to reduce the backlog. When the operator completes the task before the end of the cycle time, he or she can work ahead of schedule [16]. Production stops when the delay exceeds a critical limit. Contrary to this open station-type operating strategy, close-station production is referred to when the operator must stop the conveyor even in the event of a minor delay [17].

Figure 1. The wire harness paced assembly conveyor (often referred to as a rotary) contains assembly tables consisting of connector and clip fixtures [18].

The key idea is that in the case of modular production, the expected activity times are estimated based on the Bill of Materials (BoM) of the manufactured products. The manufacturing is modular meaning that the products $p_1, \ldots, p_{N_p}$ are built from the set of modules $m_1, \ldots, m_{N_m}$ [19]. The structures of the products are defined by a **P**-matrix (also referred to as a binary/logical matrix) consisting of N_p rows and N_m columns, and the element $p_{i,j}$ of **P** is set to one when the p_i-th type of product contains the m_j-th module (otherwise it is 0). The calculation of the theoretical activity times is estimated based on which $a_1, \ldots, a_{N_a}$ activities are needed to be performed and which $c_1, \ldots, c_{N_c}$ components should be built in at the $w_1, \ldots, w_{N_w}$ workstations. This information is represented in the logical matrix **M** that contains the activities required to produce a given product. As is shown in Table 1, the **C** matrix stores which components are built in in each activity, while the **W** matrix assigns activities to the workstations. The specific activity times and factors influencing them were determined based on expert knowledge [15] as presented in Table 2. The matrix **T** provides information on the category of the activity describing how the activities are classified into the activity types $t_1, \ldots, t_{N_t}$. The sequence of the products is represented by a π vector of the labels of the types, so $\pi(k) = p_j$ states that type product p_j started to be produced during the k-th production cycle.

To ensure fully reproducible results, only openly available information on wire harness manufacturing technologies was utilized during the development of this case study. To stimulate further research, the resultant algorithm of the developed model of the production system and the details of the products and sensor placements are publicly available on the website of the authors (https://www.abonyilab.com/soft-sensors).

Table 1. The logical matrices defined for performance monitoring.

Notation	Nodes	Description	Size
A	product (**p**) - activity (**a**)	activity required to produce a product	$N_p \times N_a$
W	activity (**a**) - workstation/machine (**w**)	workstation assigned for an activity	$N_a \times N_w$
B	product (**p**) - component/part (**c**)	component/part required to produce a product	$N_p \times N_c$
P	product(**p**) - module (**m**)	module/part family required to produce a product	$N_p \times N_m$
C	activity (**a**) - component (**c**)	component/part built in or processed in an activity	$N_a \times N_c$
M	activity (**a**) - module (**m**)	activity required to produce a module	$N_a \times N_m$
T	activity (**a**) - activity type (**t**)	category of the activity	$N_a \times N_t$
$\mathbf{S}^w$	activity (**a**) - measured time interval ($\mathbf{z}^w(k)$)	activity involved over a measured time interval	$N_a \times l_w$

Table 2. Types of activities and the related activity times according to [15]. The activity times are calculated using a direct proportionality approach, e.g., when an operator is laying four wires over one foot, proportionally to the parameter t_4, the activity time will be 1×6.9 s $+ 4 \times 4.2$ s $= 23.7$ s.

ID	Activity	Unit	Time [s]
t_1	Point-to-point wiring on chassis	Number of wires	4.6
t_2	Laying in U-channel		4.4
t_3	Laying flat cable		7.7
t_4	Laying wire(s) onto harness jig	Per wire	6.9 4.2
t_5	Laying cable connector (one end) onto harness jig	Per wire	7.4 2.3
t_6	Spot-tying onto cable and cutting		16.6
t_7	Lacing activity		1.5
t_8	Taping activity		6.8
t_9	Inserting into tube or sleeve		3.0
t_{10}	Attachment of wire terminal		22.8
t_{11}	Screw fastening of terminal		17.1
t_{12}	Screw-and-nut fastening of terminal		24.7
t_{13}	Circular connector		11.3
t_{14}	Rectangular connector		24.0
t_{15}	Clip installation		8.0
t_{16}	Visual testing		120.0

Based on the data published in [14,15], the number of types of products N_p is assumed to be 64 and defined as the combination of $N_m = 7$ modules: base module m_1, left- or right-hand drive m_2, normal/hybrid m_3, halogen/LED lights m_4, petrol/diesel engine m_5, 4 doors/5 doors m_6, and manual or automatic gearbox m_7. The number of activities/tasks N_a is defined as 654 and categorized into $N_t = 16$ types of activities. The time consumptions of these activities are approximated using a direct proportionality approach with regard to the primary activities (see Table 2). During the activities involved in the production of the base harness 115 different part families (component types, N_c) are built in (among these $C_t = 162$ terminals, $C_b = 63$ bandages, $C_c = 25$ clips, and $C_w = 89$ wires). The conveyor consists of 10 workstations (tables, N_w). For every table (workstation) one operator is assigned, therefore, $N_o = 10$.

Hereinafter, the term primary activity time denotes the estimated average period of time required for a certain type of activity to be performed, while the term local activity time refers to the time period required by a specific operator at the w-th workstation to perform the activity in question. The structure of the developed production-monitoring model is determined by the available information [15]. The proposed matrix-based mathematical formulation is beneficial as it allows the compact estimation of the individual $\hat{y}_i^w(k)$, $i = 1, \ldots, N_a$ activity times in every k cycle step (discrete time):

$$\hat{y}_i^w(k) = [\mathbf{t}_i, \mathbf{c}_i]\, \mathbf{x}^w(k), \tag{1}$$

as the time consumption of the i-th activity depends on how many elementary activities of a given type should be performed (represented as $\mathbf{t}_i$ which is the i-th row of the matrix $\mathbf{T}$), the number of built in components (the row vector $\mathbf{c}_i$ is the i-th row of the matrix $\mathbf{C}$) and the 'efficiency' of the operator $\mathbf{x}^w(k)$, which is the vector of the estimated local activity times. Therefore, the aim of our investigation is to provide a continuous local estimate of this state vector and its workstation independent $\mathbf{x}(k)$ version providing a reference value and the opportunity for the isolation of operator-independent problems.

2.2. Fixture Sensor- and Indoor Positioning System-Based Activity-Time Measurements

To measure the activity times, fixture sensors were designed as depicted in Figure 2. The fixture-based activity sensors generate timestamps when the component is inserted into the fixture. The sensors on an illustrated assembly table are shown in Figure 3.
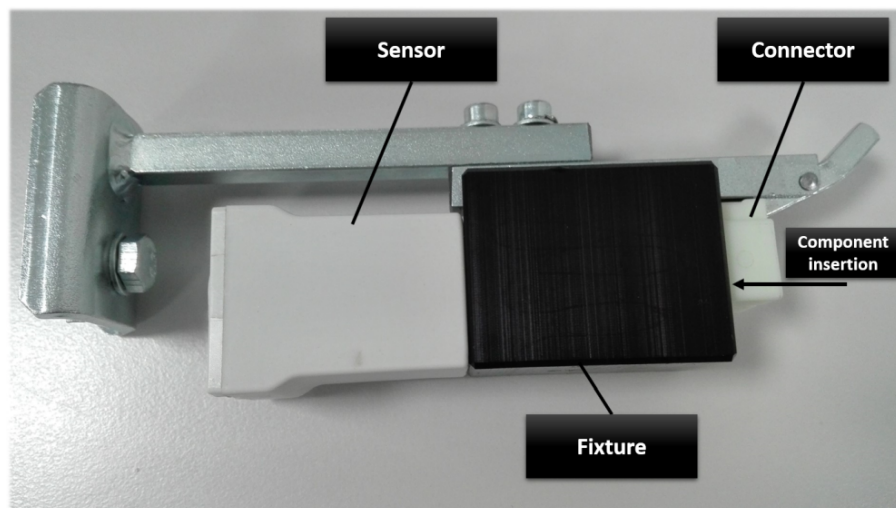
Figure 2. The designed connector fixture sends timestamps when the operator inserts a component into a fixture.

The fixtures were positioned based on how the measurable activities at the workstations are distributed. For example, the sensor f_1 sends a timestamps when the operator inserts the component c_1 which represents the starting time of the first activity a_1. Details concerning the placement of the sensors are given in Table 3.

Table 3. The placement of the sensors is defined based on the activity IDs. As can be seen in the table, not all the f_i $i = 1, \ldots, 16$ fixtures are active at every w_j $j = 1, \ldots, 10$ workstation.

Sensor ID	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}
f_1	1	79	159							
f_2	12	90	170							
f_3	21	99	175							
f_4	31	109	181							
f_5	44	121	185	226						
f_6								422	486	595
f_7								438	514	603
f_8								448	535	
f_9								451	540	615
f_{10}		132	192		275	324	373	453		
f_{11}					323	372		482		
f_{12}							419			
f_{13}										617
f_{14}										630
f_{15}									547	
f_{16}										654

The activity-dependent sequence of the timestamps recorded by the active sensors in the k-th cycle of the conveyor is represented by vector $\mathbf{s}(k) = \left[s_1(k), \ldots, s_j(k) \ldots, s_{N_s}(k) \right]^T$ which serves as the raw input of the performance-monitoring algorithm.
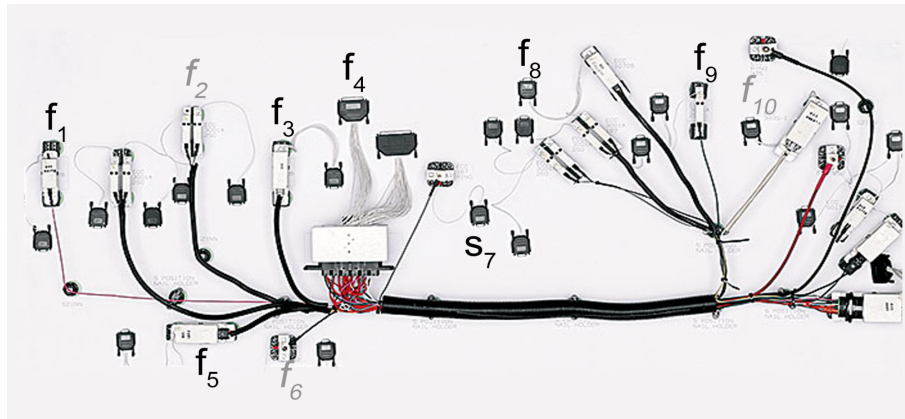
Figure 3. Illustration of the distribution of the fixtures on an assembly table. As the fixtures move according to the tables of the conveyor system, the fixtures are identically placed at every workstation. The fixtures labeled with gray text are inactive as there are no related activities at the depicted workstation.

As is shown in Figure 4, two timestamps clasp a set of activities, therefore, the $z_i^w(k) = s_{\beta(i)}^w(k) - s_{\alpha(i)}^w(k)$ difference between any two timestamps provides the sum of the activity times that are situated between the two sensors. If the timestamps $s_{\alpha(i)}^w(k)$ measures the start of the first activity at the w-th workstation, the station time of the w-th workstation can be measured as $z_i^w(k) = s_{\alpha(i)}^w(k+1) - s_{\alpha(i)}^w(k)$. Based on this concept, a set of measurements can be defined for the workstations $\mathbf{z}^w(k) = \left[z_1^w(k), \ldots, z_i^w(k), \ldots z_{l_w}^w(k) \right]^T$ which is much more interpretable and applicable information with regard to activity-time monitoring than the $\mathbf{s}(k)$ values of the raw measurements.
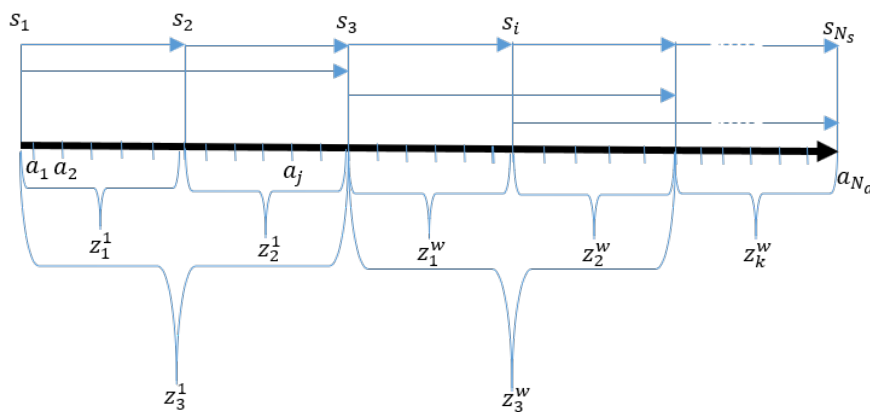

Figure 4. The concept of activity time measurements. The differences between the timestamps (s_i) define the time period required by a set of activities (a_j), the totals of which are considered as measured variables at each workstation (z_i^w), where w represents the index of the workstation.

To put $\mathbf{z}^w(k)$ into context, the information on which products are assembled at each station and the details of the activities that are assigned to the measured time interval $z_i^w(k)$ are required.

The assignment of the activities and the measured time intervals are represented by a set of logical matrices $\mathbf{S}^w$ (see Table 1). In the case of modular production the set of activities $\mathbf{q}_a = \mathbf{M}\mathbf{p}_p^T$ should be calculated based on which modules are included in the produced p-th product (represented as $\mathbf{p}_p$ which is the p-th row of the product-module matrix $\mathbf{P}$) and whose activities are required to produce the modules (such information is stored in the relation matrix $\mathbf{M}$). The activities that are assigned to the $\mathbf{z}^w(k)$-th intervals are defined by the operation $diag(\mathbf{q}_a)\mathbf{S}^w$.

$\mathbf{T}^T diag(\mathbf{q}_a)\mathbf{S}^w$ groups the activities according to activity types, while the number of components installed over a specific time interval is calculated as $\mathbf{C}^T diag(\mathbf{q}_a)\mathbf{S}^w$, which can also be grouped by activity types according to $\left(\mathbf{T}^T\mathbf{C} > 0\right)\mathbf{C}'diag(\mathbf{q}_a)\mathbf{S}^w$. Based on the proposed matrix-type representation, the estimated time intervals at the w-th workstation can be calculated as:

$$\hat{\mathbf{z}}^w(k) = \left[\mathbf{T}^T diag(\mathbf{q}_a)\mathbf{S}^w, \left(\mathbf{T}^T\mathbf{C} > 0\right)\mathbf{C}^T diag(\mathbf{q}_a)\mathbf{S}^w\right]\mathbf{x}^w(k) = \mathbf{H}^w(k)\mathbf{x}^w(k) \tag{2}$$

The model equation $\mathbf{z}^w(k) = \mathbf{H}^w(k)\mathbf{x}^w(k) + \mathbf{e}^w$ and the related measurements $\mathbf{z}^w(k)$ can be used for the continuous estimation of the vector of operator efficiencies (namely estimated local activity times), $\mathbf{x}^w(k)$, where $\mathbf{e}^w(k)$ is assumed to be a serially uncorrelated white-noise vector of observational errors with covariance matrix $\mathbf{R}^w(k)$.

As $\mathbf{H}^w(k)$ depends on the actual product, which product is produced at the w-th workstation must be tracked. For the localization of the products and identification of the status of the conveyor system, an Ultra-Wide band (UWB) Indoor Positioning System (IPS) technology with its low energy demand for transmitting information over a broad bandwidth (>500 MHz) and accuracy within the range of 30–50 cm, which is significantly better than the uncertainty of one meter that the BLE based solutions posses [11,12], was applied.

In comparison with outdoor environments, sensing location information in indoor environments requires higher precision which is a more challenging task because various objects reflect and disperse signals. Ultra-Wideband (UWB) is an emerging technology in the field of indoor positioning [20] that has shown better performance compared to others [21] even in the presence of severe multipath [22,23]. Depending on the positioning technique, the angle of arrival (AOA), the signal strength (SS), or time delay information can be used for positioning [12]. Received signal strength (RSS) UWB positioning methods also can be divided into Time of Arrival (ToA), Angle of Arrival (AoA) and Received Signal Strength (RSS)-based systems [24].

The concept of identification of the products at workstations to extract product-relevant information from the Bill of Materials (BoM) and other structured information sources are widely used to support production management [25], value stream mapping [26], and Industrial Internet of Things (IIoT)-based lifecycle management [27]. In the developed system the IPS beacons are mounted to the flat wire-harness and the raw signals of the receivers (shown in Figure 5) are processed to assign the cables to the workstations.
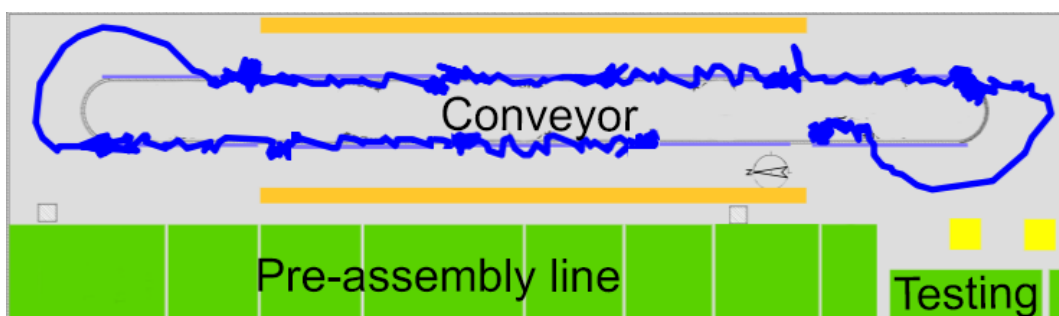

Figure 5. Illustration of how the internal positioning system (IPS) tracks a product in the conveyor at the table. The tracking is accurate and nicely depicts the rotations of the table at the edges of the conveyor.

2.3. Multi-Sensor Data Fusion-Based Recursive Estimation

Multiple sensors provide redundancy which enables the robust recursive estimation of the unmeasured primary activity times of the operators. Therefore, the estimation problem is defined as a sensor-fusion task [28]. The presented sensor fusion algorithm combines all sensory and production data such that the estimates of the activity times have less uncertainty than would be possible when

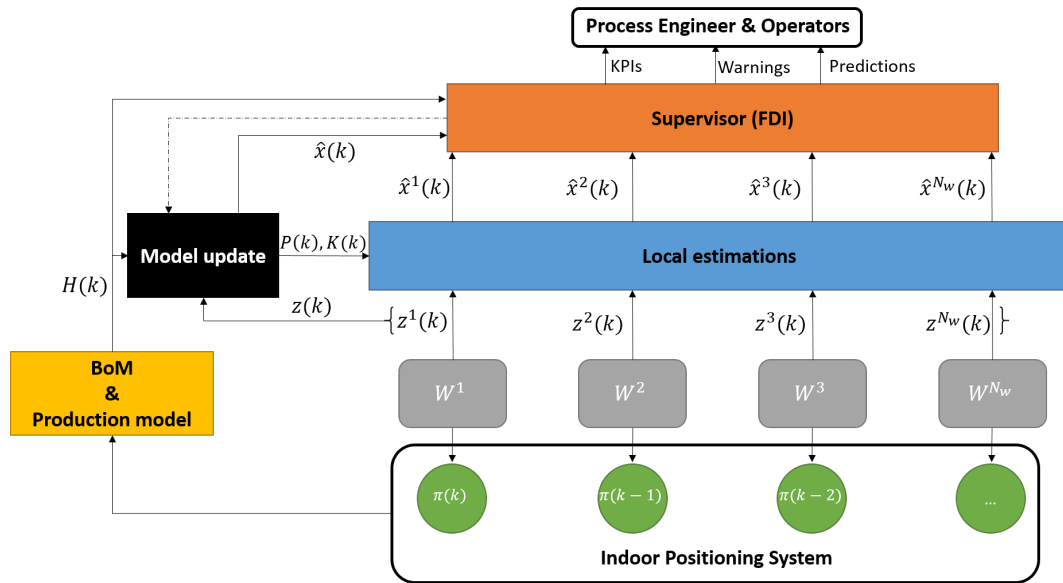these sources were used individually. The elements of the monitoring system are structured as shown in Figure 6.

Figure 6. The sensor fusion-based architecture of the proposed monitoring system.

The fusion center receives and synchronizes all the $\mathbf{z}^w(k), w = 1, \ldots, N_w$ measured time intervals and the related $\mathbf{H}^w(k), w = 1, \ldots, N_w$ time-variable regressors, which means all data collected from the workstations are time-stamped and arranged according to k-th cycle of the conveyor:

$$\mathbf{z}(k) = \begin{bmatrix} \mathbf{z}^1(k) \\ \vdots \\ \mathbf{z}^{N_w}(k) \end{bmatrix}, \mathbf{H}(k) = \begin{bmatrix} \mathbf{H}^1(k) \\ \vdots \\ \mathbf{H}^{N_w}(k) \end{bmatrix}, \tag{3}$$

The linear structure of the developed production-monitoring model (see Equation (1)) is adequate for the studied problem as the time consumption of the activities linearly depend on how many elementary activities should be performed and what is the number of the built in components [15]. When a linear sensor-fusion model is assumed, the previously presented linear time-variant model can be represented as

$$\mathbf{z}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{e}(k), \tag{4}$$

where the $\mathbf{e}(k)$ noise vector of the fused observations consists of the $\mathbf{e}^w(k)$ serially uncorrelated white-noise vectors of observational errors at the workstations, $\mathbf{e}(k) = \left[\left(\mathbf{e}^1(k) \right)^T, \ldots, \left(\mathbf{e}^{N_w}(k) \right)^T \right]^T$.

When the observation errors of the workstations are assumed to be independent, the covariance of the $\mathbf{e}(k)$ noise vector is a block diagonal matrix defined as $\mathbf{R} = diag\left(\mathbf{R}^1, \ldots, \mathbf{R}^{N_w} \right)$.

The central estimation enhances the confidence of the nominal model which improves the performance of fault detection based on the reconciliation of the local measurements [29].

Based on $k = 1, \ldots, N$ synchronized $\mathbf{z}(k)$ and $\mathbf{H}(k)$ observations the objective function of the central estimation problem can be formalized as:

$$\hat{\mathbf{x}}(N) = \arg\min_{\mathbf{x}} V_N(\mathbf{x}) \ V_N(\mathbf{x}) = \frac{1}{N} \sum_{k=1}^{N} [\mathbf{z}(k) - \mathbf{H}(k)\mathbf{x}]^T \mathbf{Q} [\mathbf{z}(k) - \mathbf{H}(k)\mathbf{x}] \tag{5}$$

When the positive-definite weighting matrix $\mathbf{Q}$ is defined as $\mathbf{Q} = (\mathbf{R})^{-1}$, the estimation is equivalent to the maximum-likelihood cost function [30].

The covariance matrix of the estimation error $\tilde{\mathbf{x}}(k) = \hat{\mathbf{x}}(k) - \mathbf{x}(k)$ is:

$$E\left(\tilde{\mathbf{x}}(N)\tilde{\mathbf{x}}^T(N)\right) = \mathbf{P}^*(N) = \left[\sum_{k=1}^{N} \mathbf{H}^T(k)\mathbf{R}^{-1}\mathbf{H}(k)\right]^{-1} \tag{6}$$

The recursive estimation of the primary activity times $\mathbf{x}(k)$ is similar to the state estimation algorithm which assumes the following Gauss-Markov (GM) process:

$$\mathbf{x}(k) = \mathbf{A}^*(k)\mathbf{x}(k-1) + \eta(k-1), \ \eta(k) = \mathcal{N}(\mathbf{0}, \mathbf{Q}_x) \tag{7}$$
$$\mathbf{z}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{e}(k), \ \mathbf{e}(k) = \mathcal{N}(\mathbf{0}, \mathbf{R}) \tag{8}$$

where $\eta(k)$ noise vector and its $\mathbf{Q}_x$ covariance matrix represents the uncertainty of the unknown and time-varying parameters and $\mathbf{A}^*(k)$ stands for the state transition matrix of this random process.

The recursive estimation consists of prediction and correction steps as follows.

At the prediction step the state vector and its covariance matrix is calculated based on information available at the $k-1$ time instant:

$$\hat{\mathbf{x}}(kk-1) = \hat{\mathbf{x}}(k-1) \tag{9}$$
$$\mathbf{P}^*(kk-1) = \mathbf{P}^*(k-1) + \mathbf{Q}_x \tag{10}$$

The correction step utilizes the measured $\mathbf{z}(k)$ measurements at the correction the estimated state variables by the $\mathbf{e}(k) = [\mathbf{z}(k) - \mathbf{H}(k)\hat{\mathbf{x}}(kk-1)]$ prediction error, with the $\mathbf{K}(k)$ time-varying Kalman gain updated based on the refreshed $\mathbf{P}^*(k)$ covariance matrix:

$$\hat{\mathbf{x}}(k) = \hat{\mathbf{x}}(kk-1) + \mathbf{K}(k)\left[\mathbf{z}(k) - \mathbf{H}(k)\hat{\mathbf{x}}(kk-1)\right] \tag{11}$$
$$\mathbf{K}(k) = \mathbf{P}^*(kk-1)\mathbf{H}^T(k)\left[\mathbf{R} + \mathbf{H}(k)\mathbf{P}^*(kk-1)\mathbf{H}^{*T}(k)\right]^{-1} \tag{12}$$
$$\mathbf{P}^*(k) = \mathbf{P}^*(kk-1) - \mathbf{K}(k)\mathbf{H}(k)\mathbf{P}^*(kk-1) \tag{13}$$

2.4. Local Estimation and Monitoring of the Primary Activity Times

To constrain the model parameters to lie within a reliable region and incorporate important *a priori* knowledge of the activity times, the estimated parameters were optimally projected on to the set of linear constraints by quadratic programming [13].

The local (operator-related) projection of the unconstrained estimate $\hat{\mathbf{x}}(k)$ can be considered as a quadratic programming problem:

$$\hat{\mathbf{x}^w}(k) = arg \min_{\mathbf{x}(k)}[\mathbf{x}(k) - \hat{\mathbf{x}}(k)]^T\mathbf{Q}_p[\mathbf{x}(k) - \hat{\mathbf{x}}(k)] \tag{14}$$

subject to:

$$\mathbf{A}_e^w(k)\mathbf{x}(k) = \mathbf{b}_e^w(k) \tag{15}$$
$$\mathbf{L}^w\mathbf{x}(k) \leq \mathbf{c}^w \tag{16}$$

$$\hat{\mathbf{x}}(k)^c = \hat{\mathbf{x}}(k) - \mathbf{P}^*(k)\mathbf{H}_j^T\mu_j - \mathbf{P}^*(k)\mathbf{L}_j^T\lambda_j \tag{17}$$

where $\hat{\mathbf{x}}(k)$ denotes the unconstrained solution, $\hat{\mathbf{x}}(k)^c$ denote the constrained solution, $\mathbf{A}_e^w(k)$ and $\mathbf{b}_e^w(k)$ define the linear equality constraints, while $\mathbf{L}^w(k)$ and $\mathbf{c}^w(k)$ represent the linear inequalities. μ_j and λ_j are vectors of Lagrange multipliers associated with equality and inequality constraints. This formulation ensures the optimal (least squares correction) when $\mathbf{Q}_p = (\mathbf{P}^*(k))^{-1}$. When $\mathbf{Q}_p$ denotes the identity matrix an orthogonal projection is obtained. Assuming the constraints are true, parameter bias can never be increased [13].

The following section demonstrates how the estimated and expected primary activity times are used for production monitoring.

3. Wire Harness Case Study

3.1. Online Monitoring of Operator Performance

To validate the reliability of the proposed model, the distribution of the activity times collected from real production lines was studied. As is illustrated in Figure 7 the distribution of the assembly times can be broken down into several Gaussian-type distributions.
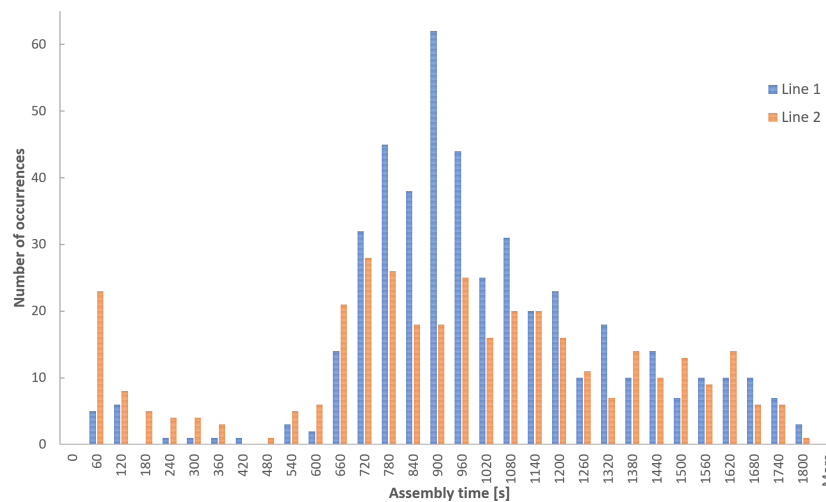

Figure 7. The histogram of measured processing times in two different conveyors (production lines). The histograms indicate that the distribution of the sensor-delivered processing times can be decomposed into normal distribution functions according to different products.

The identifiability of the model is determined by the rank of the covariance matrix $\mathbf{P}^*(N)$. When the rank is smaller than the number of measurements (which occurs when the individual performance of operators is estimated at a specific workstation) only a subset of the parameters is identifiable.

The information content of the available data can be evaluated based on the eigenvalues or determinant of the covariance matrix $\mathbf{P}^*(N)$. The tools of D-optimal experimental design that tries to maximize the determinant of $\mathbf{F}^*(N)$ which is identical to the minimization of the determinant of $\mathbf{P}^*(N)$ where utilized.

$$\mathbf{F}^*(N) = (\mathbf{P}^*(N))^{-1} = \sum_{k=1}^{N} \mathbf{H}^T(k)\mathbf{R}^{-1}\mathbf{H}(k) \tag{18}$$

When only one product is produced, $\mathbf{H}(k)$ does not change in terms of time. In this case, the set of the identifiable parameters for a given product can be determined by the QR decomposition of $\mathbf{H}(k)$ (or $\mathbf{H}^w(k)$ when a local estimation is needed). When different products are produced, the variation in $\mathbf{H}(k)$ significantly increases the available information, so the optimization of the production sequence can highly influence the identifiability of the model and confidence in the parameters ($\mathbf{P}^*(N)$, $\pi(k)$).

The production of 1000 products was studied. The production sequence contained all 64 types of products with an average batch size of 10 products/batch. The rank of the covariance matrix $\mathbf{F}^*(N)$ was identical to the size of $\hat{\mathbf{x}}(k)$, so all activities could be monitored (see Figure 8).

When the raw material, design or the processing of a component in a cost-cutting or quality-improvement project is changed by the supplier, this change may influence the activity times of the operators. Such operator-independent loss in performance can occur when a shorter length

of wire increases the time required to lay and arrange the cables. In this case study, such effects are monitored. In the studied case, the new wires between the c_{87} and c_8 components are a bit shorter than specified. The component c_{87} (seal on the terminal) has an impact on the t_{10} type of activity in the module m_4 which increases the related primary activity time ($x_{10}(k)$) by 15% at the 200th product, while the component c_8 (the shorter wire) has an impact on the activity type t_5 in the module m_2, which increases the related $x_5(k)$ state variable by 20% after the 300th product. In this illustrative scenario the quality inspection time decreases after the 500th product.

Figure 8. Estimated primary activity times with their $p = 0.01$ confidence intervals (represented by dashed lines). The figure illustrates that the algorithm is able to track the changes in the $x_{10}(k)$, $x_5(k)$ and $x_{16}(k)$ activity times after the 200th, 300th and 500th product, respectively. The bold lines represent the constrained parameter estimates.

As Figure 8 illustrates, the proposed system is able to track the slowed and fastened activities. The benefit of the proposed constrained algorithm is clearly visible, the estimated variables converge faster and are always reliable.

The means of detecting individual losses in operator performance losses and sensor faults (due to delayed registration and IIoT communication) were also studied.

In terms of fault detection, the prediction error used in Equation (5) can be used as generates an interpretable and easily traceable univariate time series that reflects the global performance of the model.

The global performance of the model is reflected by

$$e_q(k) = [\mathbf{z}(k) - \mathbf{H}(k)\mathbf{x}]^T \mathbf{Q} [\mathbf{z}(k) - \mathbf{H}(k)\mathbf{x}] \,, \tag{19}$$

while the local, workstation related fault detection should be based on the local observations:

$$e_q^w(k) = [\mathbf{z}^w(k) - \mathbf{H}^w(k)\mathbf{x}]^T \mathbf{Q}^w [\mathbf{z}^w(k) - \mathbf{H}^w(k)\mathbf{x}^w] \,, \tag{20}$$

where $\mathbf{Q}^w$ represents the wth block matrix of $\mathbf{Q}$.

Based on the analysis with regard to the rank of the $\mathbf{H}^w(k)$ matrices, the observable sets of activities were determined. As is illustrated in Figure 9, at the $w = 2$ workstation the time consumption of six primary activities are observable. The proposed algorithm was able not only to detect operator-dependent problems (of the 250th product) related to these activities, but by monitoring the $e_q(k)$ it was possible to determine when sensor faults occurred (see the bottom of the figure). The parameters of the gross error detection algorithm can be fine-tuned by Monte Carlo simulation and detailed analysis of the distribution of the modeling error [31,32] (the demonstration of the applicability of these techniques in this problem is out of the scope side this paper).

As is illustrated in Figure 10, the calculations above can be used to estimate the expectable operation times for all workstations, check how well the process is balanced and how the complexity of the product influences the workloads of the workstations. With the help of this model the effect of the changes in the activity time can be immediately calculated on the tack-time and the effectiveness of the operators. The presented example demonstrated that in the event of good estimates with regard to the duration of the primary activities and with the help of the IIoT-based fusion of product-relevant information, real-time data for Overall Equipment Effectiveness (OEE) calculations can be provided.
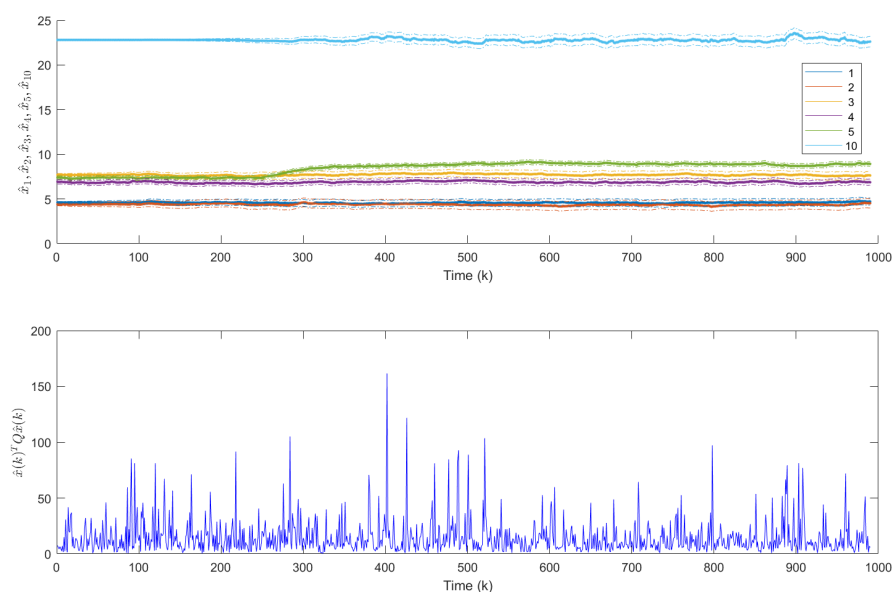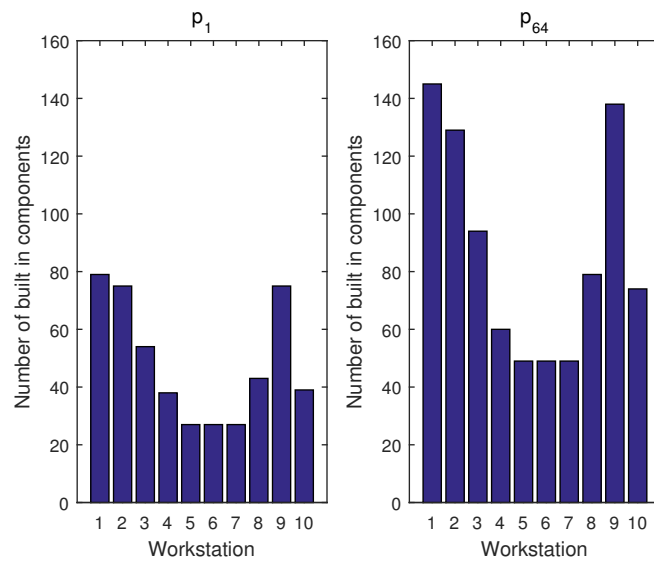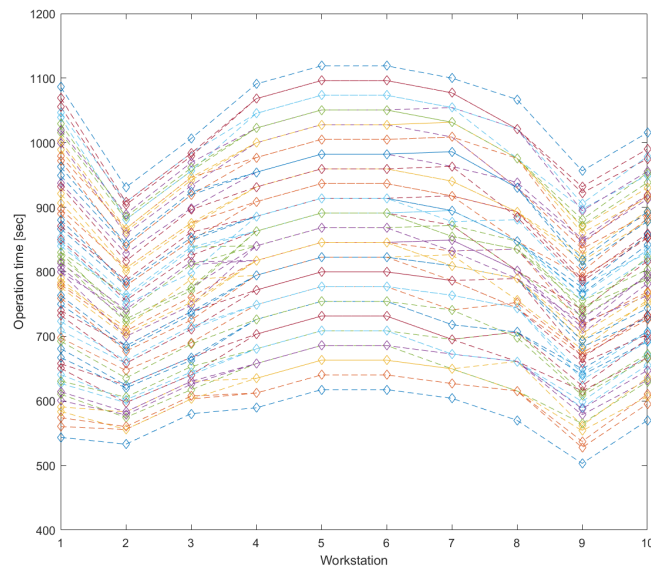

Figure 9. Fault-detection performance at the 2nd workstation. The upper figure illustrates that the algorithm is able to detect operator-dependent problems (after the 250th product). After this change the related to these activities.

(**a**) The number of the built-in components at a given workstation. The figure shows how the workload differs during the production of the base module (p_1) and the most complex product (p_{64}).

(**b**) The variability of the station times during the production of the 64 product. The figure illustrates how the production line is balanced and how the complexity of different products influences the station times.

Figure 10. The workloads (number of activities, built-in components and total activity times) can be easily calculated based on the proposed model. The OEE of the production line can be monitored on-line based on the recursively estimated activity times.

3.2. Targeting Model-Based Workload Analysis

The monitoring of the activity of operators is based on the comparison of the measured activity and station times with the estimates of targeting models whose parameters are identified by the proposed

estimation algorithms. Targeting models are widely used in process performance management. Precedent-based targeting models are used when the expected performance can be deduced from previous operations (e.g., a day or month before). One weakness of this procedure is that it assumes that conditions were comparable over the two time intervals. A more problematic issue is what happens when a significant change occurs in the process and product. For such applications precedent-based targeting models can be over simplistic. Activity-based targeting is particularly appropriate when the clear drivers of performance are known [33].

Automatic monitoring and targeting schemes attempt to compare performances very short time intervals (e.g., minutes) which is ideal for fault detection. The most important key performance indicators (KPIs) of the production system are the station times which reflect how well the production line is balanced. The concept of calculating the station time is depicted in Figure 11. The balancing of a modular production system is a challenging industrial problem due to the great diversity of products [34]. As the station times are the functions of the manufactured products, which product is assembled on a given workstation must be followed. The calculation of the station time is similar to the calculation of the estimated sum of activity times between two fixture sensors (Equation (2)), namely the difference between the appropriate timestamps recorded by the fixture sensors:

$$\mathbf{y}^w(k) = \left[\mathbf{T}^T \mathbf{q}_w \, , \, \left(\mathbf{T}^T \mathbf{C} > 0 \right) \mathbf{C}^T \mathbf{q}_w \right] \mathbf{x}^w(k) \tag{21}$$

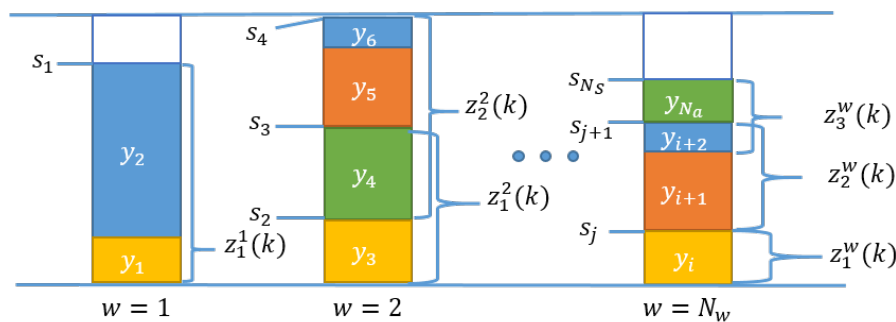where $\mathbf{q}_w = diag\left(\mathbf{M}\mathbf{p}'_p(k) \right) \mathbf{W}$.

Figure 11. Demonstration of the distribution of the activity times y_i, the station times, and the timestamps s_j of the fixture sensors; the measured time interval z_l^w; and the w-th workstation. The height of the bars coloured illustrates the activity time of the specific activity, while the coloured tag of the individual activities provides information on the type of the specific activity. This representation shows how the related activity times are summed up and form the station time and highlights that the longest station time is the cycle time of the production system, which is represented by the vertical blue line at the top of the figure.

4. Conclusions

Human-in-the-loop cyber-physical production systems are transforming the industrial workforce. Due to the enormous variability and complexity of products, the tracing of hundreds of activity times on production lines is a critical problem. To handle this problem a software-sensor-based activity-time and performance measurement system was proposed. To ensure a real-time connection between operator performance and varying degrees of product complexity fixture sensors were utilized and designed and an indoor positioning system used to merge this multi-sensor data with product-relevant information.

The presented sensor fusion algorithm combines all sensory and production data such that the estimates of the activity times have less uncertainty than would be possible when these sources were used individually. The estimation of the activity times is based on a linear-in-parameters model. The linear structure of the developed production-monitoring model is adequate as the time

consumption of the activities linearly depend on how many primary activities should be performed and what is the number of the built-in components.

The number of parameters of activity time estimation models is comparable to the number the number of measurements, the identifiability of the parameters of the model has to be carefully analyzed. For this purpose, we studied the Fisher information/covariance matrix of the estimation problem. The identifiability of the model and the information content of the available data can be evaluated based on the rank, the eigenvalues and the determinant of the covariance matrix. When the rank is smaller than the number of measurements (which occurs when the individual performance of operators is estimated at a specific workstation), only a subset of the parameters is identifiable. As the placement of the sensors significantly influences the identifiability of the parameters, tools of D-optimal experimental design can be used to optimize the proposed system.

The determination of the optimal number of sensors and features has crucial importance as redundant sensors can generate correlated features which decrease the efficiency of the algorithm. The analysis of the eigenvalues of the covariance matrix can highlight these negative effects. As this analysis is identical to Principal Component Analysis (PCA) of the multisensor data, the proposed methodology can utilize the reduced and transformed uncorrelated features, which results in a Principal Regression-based process monitoring algorithm. The second approach of avoiding correlated features is the application of feature selection algorithms that should be based on the previously discussed experimental design optimization task.

As the estimation problem can be ill-conditioned and poor raw sensor data can result in unrealistic parameter estimates, constraints were introduced into the parameter-estimation algorithm to increase the robustness of the software sensor.

The proposed model-based performance monitoring system tracks the recursively estimated parameters of the activity-time estimation models, while the sensor-relevant fault detection functionalities are based on the modeling errors which can be evaluated by classical residual-based fault detection algorithms.

The applicability of the proposed methodology is demonstrated on a well-documented benchmark problem of a wire harness manufacturing process. The presented example demonstrated the benefits of multiple sensors as they provide redundancy which enables the robust recursive estimation of the unmeasured primary activity times. The fully reproducible and realistic simulation study also confirmed the efficiency of the proposed constrained estimation algorithm regarding fast convergence and giving reliable estimates.

The results illustrate that indoor positioning system-based integration of product-relevant information and sensor signals and can be efficiently utilized to design on-line performance management systems.

The developed benchmark problem can be used to study fault detection and sensor placement algorithms which is the objective of our further research.

Author Contributions: T. Ruppert and J. Abonyi developed the methodology and developed the software; T. Ruppert prepared the original draft; J. Abonyi reviewed and edited the manuscript.

Notations and Abbreviations

CPPS	Cyber-physical production system
IIoT	Industrial Internet of Things
RFID	Radio Frequency IDentification
BoM	Bill of Materials
IPS	Indoor positioning system
UWB	Ultra-Wide Band
BLE	Bluetooth Low Energy
MSDF	Multi-sensor data fusion
AoA	Angle of Arrival
SS	Signal Strength
RSS	Received Signal Strength
ToA	Time of Arrival
$p_1, \ldots, p_{N_p}$	products
$m_1, \ldots, m_{N_m}$	modules
$a_1, \ldots, a_{N_a}$	activities
$c_1, \ldots, c_{N_c}$	components
$w_1, \ldots, w_{N_w}$	workstations
$t_1, \ldots, t_{N_t}$	activity types
$\mathbf{A}$	$(N_p \times N_a)$ activities required to produce a product
$\mathbf{W}$	$(N_a \times N_w)$ workstation assigned for an activity
$\mathbf{B}$	$(N_p \times N_c)$ component/part required to produce a product
$\mathbf{P}$	$(N_p \times N_p)$ module/part family required to produce a product
$\mathbf{C}$	$(N_a \times N_c)$ component/part built in or processed in an activity
$\mathbf{M}$	$(N_a \times N_m)$ activity required to produce a module
$\mathbf{T}$	$(N_a \times N_t)$ category of the activity
$\mathbf{S}^w$	$(N_a \times l_w)$ activity involved over a measured time interval
k	index of the production cycle (discrete time)
$\hat{y}_i^w(k)$	estimation of the individual activity times for work station w in the kth production cycle
$\mathbf{x}^w(k)$	'efficiency' of the operator, the vector of the estimated local activity times
$\mathbf{x}(k)$	workstation-independent version of $\mathbf{x}^w(k)$
$\mathbf{s}(k)$	sequence of the timestamps recorded by the active fixture sensors
$\mathbf{z}^w(k)$	vector of the sum of the activity times that are situated between the two sensors
α	index of the first sensor of a fixture-sensor pair
β	index of the second sensor of a fixture-sensor pair
$\mathbf{q}_a$	the set of activities required to produce a specific product
$\mathbf{e}^w$	serially uncorrelated white-noise vector of observational errors
$\mathbf{R}^w(k)$	covariance matrix of observational errors
$\mathbf{H}(k)$	time-variable regressors representing the number of activities and built in components
$\mathbf{e}(k)$	the set of the serially uncorrelated white-noise vector of observational errors of the workstations
$\hat{\mathbf{x}}(N)$	estimation error
$\mathbf{Q}$	positive-definite weighting matrix defined as $\mathbf{Q} = (\mathbf{R})^{-1}$
$\mathbf{P}^*$	inverse of the parameter covariance matrix
$\mathbf{A}^*(k)$	State-transition matrix in the Kalman filter represented estimation problem (in our case an identity matrix)
$\mathbf{K}(k)$	Gain of the Kalman filter/recursive estimator
$\mathbf{L}^w, \mathbf{c}^w$	Representation of the linear inequality constraints
$\mathbf{A}_e^w, \mathbf{b}_e^w$	Representation of the linear equality constraints
μ_j	vector of Lagrange multiplier associated with equality
λ_j	vector of Lagrange multiplier associated with inequality constraints

References

1. Romero, D.; Stahre, J.; Wuest, T.; Noran, O.; Bernus, P.; Fast-Berglund, Å.; Gorecky, D. Towards an operator 4.0 typology: A human-centric perspective on the fourth industrial revolution technologies. In Proceedings of the International Conference on Computers and Industrial Engineering (CIE46), Tianjin, China, 29–31 October 2016.
2. Wang, S.; Wan, J.; Li, D.; Zhang, C. Implementing smart factory of industrie 4.0: An outlook. *Int. J. Distrib. Sens. Netw.* **2016**, *12*, 1–12. [CrossRef]
3. Bidanda, B.; Ariyawongrat, P.; Needy, K.L.; Norman, B.A.; Tharmmaphornphilas, W. Human related issues in manufacturing cell design, implementation, and operation: A review and survey. *Comput. Ind. Eng.* **2005**, *48*, 507–523. [CrossRef]
4. Roitberg, A.; Perzylo, A.; Somani, N.; Giuliani, M.; Rickert, M.; Knoll, A. Human activity recognition in the context of industrial human-robot interaction. In Proceedings of the Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific, Siem Reap, Cambodia, 9–12 December 2014; pp. 1–10.
5. Krahnstoever, N.; Rittscher, J.; Tu, P.; Chean, K.; Tomlinson, T. Activity recognition using visual tracking and RFID. In Proceedings of the 2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05), Breckenridge, CO, USA, 5–7 January 2005; Volume 1, pp. 494–500.
6. Stiefmeier, T.; Roggen, D.; Ogris, G.; Lukowicz, P.; Tröster, G. Wearable activity tracking in car manufacturing. *IEEE Pervasive Comput.* **2008**, *7*, 1–7. [CrossRef]
7. Koskimaki, H.; Huikari, V.; Siirtola, P.; Laurinen, P.; Roning, J. Activity recognition using a wrist-worn inertial measurement unit: A case study for industrial assembly lines. In Proceedings of the 2009 17th Mediterranean Conference on Control and Automation, Thessaloniki, Greece, 24–26 June 2009; pp. 401–405.
8. Ward, J.A.; Lukowicz, P.; Troster, G.; Starner, T.E. Activity recognition of assembly tasks using body-worn microphones and accelerometers. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 1553–1567. [CrossRef] [PubMed]
9. Huikari, V.; Koskimäki, H.; Siirtola, P.; Röning, J. User-independent activity recognition for industrial assembly lines-feature vs. instance selection. In Proceedings of the 5th International Conference on Pervasive Computing and Applications, Maribor, Slovenia, 1–3 December 2010; pp. 307–312.
10. Voulodimos, A.S.; Doulamis, N.D.; Kosmopoulos, D.I.; Varvarigou, T.A. Improving multi-camera activity recognition by employing neural network based readjustment. *Appl. Artif. Intell.* **2012**, *26*, 97–118. [CrossRef]
11. Faragher, R.; Faragher, R. An Analysis of the Accuracy of Bluetooth Low Energy for Indoor Positioning Applications. In Proceedings of the 27th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+'14), Tampa, FL, USA, 8–12 September 2014; Volume 10, pp. 201–210.
12. Gezici, S.; Tian, Z.; Giannakis, G.B.; Kobayashi, H.; Molisch, A.F.; Poor, H.V.; Sahinoglu, Z. Localization via ultra-wideband radios: A look at positioning aspects of future sensor networks. *IEEE Signal Proc. Mag.* **2005**, *22*, 70–84. [CrossRef]
13. Timmons, W.; Chizeck, H.; Casas, F.; Chankong, V.; Katona, P. Parameter-constrained adaptive control. *Ind. Eng. Chem. Res.* **1997**, *36*, 4894–4905. [CrossRef]
14. Ong, N.S. Activity-based cost tables to support wire harness design. *Int. J. Prod. Econ.* **1993**, *29*, 271–289. [CrossRef]
15. Ong, N.S.; Boothroyd, G. Assembly times for electrical connections and wire harnesses. *Int. J. Adv. Manuf. Technol.* **1991**, *6*, 155–179. [CrossRef]
16. Bukchin, J.; Tzur, M. Design of flexible assembly line to minimize equipment cost. *IIE Trans.* **2000**, *32*, 585–598. [CrossRef]
17. Sarker, B.R.; Pan, H. Designing a Mixed-Model, Open-Station Assembly Line Using Mixed-Integer Programming. *J. Oper. Res. Soc.* **2001**, *52*, 545–558. [CrossRef]
18. Assembly Line Conveyor Systems. 2015. Available online: https://www.pacline.com/photos/photos-by-solution/assembly-line-conveyors/ (accessed on 1 June 2018).
19. Da Cunha, C.; Agard, B.; Kusiak, A. Design for Cost: Module-Based Mass Customization. *IEEE Trans. Autom. Sci. Eng.* **2007**, *4*, 350–359. [CrossRef]

20. Liu, H.; Darabi, H.; Banerjee, P.; Liu, J. Survey of wireless indoor positioning techniques and systems. *IEEE Trans. Syst. Man Cybern. Part C* **2007**, *37*, 1067–1080. [CrossRef]

21. Alarifi, A.; Al-Salman, A.; Alsaleh, M.; Alnafessah, A.; Al-Hadhrami, S.; Al-Ammar, M.A.; Al-Khalifa, H.S. Ultra wideband indoor positioning technologies: Analysis and recent advances. *Sensors* **2016**, *16*, 707. [CrossRef] [PubMed]

22. Ingram, S.; Harmer, D.; Quinlan, M. Ultrawideband indoor positioning systems and their use in emergencies. In Proceedings of the PLANS 2004. Position Location and Navigation Symposium (IEEE Cat. No. 04CH37556), Monterey, CA, USA, 26–29 April 2004; pp. 706–715.

23. Lai, Z.; Joshi, H.; Goeckel, D.; Gupta, D.; Gupta, D.; Mathew, A. Performance of UWB systems in the presence of severe multipath and narrowband interference. In Proceedings of the 2008 IEEE International Conference on Ultra-Wideband, Hannover, Germany, 10–12 September 2008; Volume 1, pp. 85–88.

24. Gigl, T.; Janssen, G.J.; Dizdarevic, V.; Witrisal, K.; Irahhauten, Z. Analysis of a UWB indoor positioning system based on received signal strength. In Proceedings of the 2007 4th Workshop on Positioning, Navigation and Communication, Hannover, Germany, 22 March 2007; pp. 97–101.

25. Cunningham, M.; Higgins, P.; Browne, J. A Decision Support Tool for Planning Bills of Material. *Prod. Plan. Control* **1996**, *7*, 312–328. [CrossRef]

26. Wang, C.; Quesada-Pineda, H.; Kline, D.E.; Buehlmann, U. Using Value Stream Mapping to Analyze an Upholstery Furniture Engineering Process. *For. Prod. J.* **2011**, *61*, 411–421. [CrossRef]

27. Tao, F.; Zuo, Y.; Xu, L.D.; Lv, L.; Zhang, L. Internet of Things and BOM-Based life cycle assessment of energy-saving and emission-reduction of products. *IEEE Trans. Ind. Inform.* **2014**, *10*, 1252–1261. [CrossRef]

28. Raol, J.R. *Multi-Sensor Data Fusion with MATLAB®*; CRC Press: Boca Raton, FL, USA, 2009.

29. Li, X.R.; Zhu, Y.; Wang, J.; Han, C. Optimal linear estimation fusion. I. Unified fusion rules. *IEEE Trans. Inform. Theory* **2003**, *49*, 2192–2208. [CrossRef]

30. Young, P.C. *Recursive Estimation and Time-Series Analysis: An Introduction*; Springer Science & Business Media: Berlin, Germany, 2012.

31. Crowe, C.M. Recursive identification of gross errors in linear data reconciliation. *AIChE J.* **1988**, *34*, 541–550. [CrossRef]

32. Tong, H.; Crowe, C.M. Detection of gross erros in data reconciliation by principal component analysis. *AIChE J.* **1995**, *41*, 1712–1722. [CrossRef]

33. Moss, K.J. *Energy Management Building*; Taylor & Francis: Didcot, UK, 2006.

34. Pil, F.K.; Holweg, M. Linking Product Variety to Order-Fulfillment Strategies. *Interfaces* **2004**, *34*, 394–403.10.1287/inte.1040.0092. [CrossRef]