Journal of Engineering, Project, and Production Management
2023, 13(2), 125-137

# Correlation of Construction Workers' Movement and Direct Work Rates

Søren Wandahl[1], Cristina Toca Pérez [2], Stephanie Salling[3], and Hasse Højgaard Neve[4]
[1]Professor, Dept. of Civil & Arch. Eng., Aarhus University, Denmark, E-mail: swa@cae.au.dk (corresponding author).
[2]Postdoctoral researcher, Dept. of Civil & Arch. Eng., Aarhus University, Denmark, E-mail: cristina.toca.perez@cae.au.dk
[3]Research Assistant, Dept. of Civil & Arch. Eng., Aarhus University, Denmark, E-mail: stsa@cae.au.dk
[4]Consultat, PricewaterhouseCoopers, Denmark, E-mail: hasse.hojgaard.neve@pwc.com

Project Management
Received August 11, 2022; revised November 8, 2022; January 22, 2023; accepted January 22, 2023
Available online April 10, 2023

---

**Abstract:** The Work Sampling (WS) technique, used worldwide to understand how workers spend their time, represents a time-consuming and costly activity. Therefore, several researchers work on different approaches to automate the data collection using sensor-based and vision-based technologies. The challenge of all the sensor-based approaches is that they do not provide the share of time in different work categories. The lack of knowledge on a possible correlation between Direct Work and, e.g., presence, location, or worker movement represents a gap in the current body of knowledge. Thus, this research aims to understand the correlation between Direct Work as the independent predictor variable; and Movement as the dependent response variable**.** The authors used the data gathered through the application of WS in five case studies on building renovation projects in Denmark. To explain this correlation. The authors selected a combination of four quantitative techniques: (1) curve estimation; (2) linear regression; (3) ANOVA analysis; and (4) t-test. The correlation of the result is discussed considering three assumptions: (1) the structure of the day; (2) global vs. local; and (3) Movement vs. Transporting and Walking. The result shows a significant correlation between Direct Work and Movement with an average $R^2$ of 0.328. This is considered acceptable predictability taking the socio-technical system aspect of a construction site into account.

**Keywords**: Construction, work sampling, walking, transporting, efficiency.


DOI 10.32738/JEPPM-2023-0013

---

## 1. Introduction

Work Sampling (WS) was introduced in the early nineteen hundreds as a technique to observe workers' efficiency and productivity (Barnes, 1968). WS is deployed to determine how workers spend their time on different work activities, and it became popular, among other reasons, due to its straightforward application. The theory of WS is based on the laws of probability, which indicate that independent observations made at repeated random times will have the same distribution. Thus, random observations can be translated into percentages of time spent in activity categories (Barnes, 1968).

Over the years, the technique has been employed by practitioners and researchers for several different purposes: (a) to measure and conceptualize flow and workflow (Kalsaas, 2011; Wernicke et al., 2017); (b) to identify the share of time spent on a single activity of the same construction process on different job sites, e.g., transport (Pérez et al., 2015); (c) to provide insight for comparing the average productive workforce utilization to respective work processes in various projects (Picard, 2002); (d) to measure labor efficiency and inefficiency (Neve et al., 2021; (e) to set up a baseline measure for improvement and to serve as a challenge to management and the workers (Neve and Wandahl, 2018); (f) to understand the evolution of the share of time spent in different work categories along the years (Wandahl et al., 2021), among others.

In most cases, researchers and practitioners focused on understanding the share of time spent in different work categories. The WS categories have changed over time due to interpretation and application discrepancies. Before 1985, WS studies adopted a two-category classification of direct and non-direct work. This, to some extent, reflects Ohno's (1988) understanding of work as divided into Waste Work (WW) and Value-Added Work (VAW). However, Ohno, considered the father of the Toyota Production System, which inspired Lean Manufacturing philosophy, concluded that the VAW category must be further divided into Direct Work (DW) and Non-Value-Added-Work (NVAW). NVAW does not add value but is needed under the existing work conditions, e.g., transportation of material. The DW category is generally understood as the amount of direct, physical, and output-producing work. It can be seen as the time a worker spends producing tangible output, e.g., square meters of bricks installed (Choy and Ruwanpura, 2006). Most WS studies generally agree on this definition of DW (Wandahl et al., 2021). However, for the NVAW category, a considerable inconsistency in concept and terminology appears. Some studies categorize all NVAW as WW, while other studies have a more detailed view of NVAW as several subcategories like preparatory work, transportation, etc. Generally speaking, NVAW in WS is referred to as Indirect Work (IW), resulting in WS having three categories of time, namely DW, IW, and WW (Wandahl et al., 2021).

The non-direct work or unproductive work category is the opposite of DW and has traditionally been quite inconsistent and included everything besides DW, such as supportive work (e.g., transporting bricks to the final destination by hand) and waiting time (e.g., waiting to receive bricks in the place of execution). The non-work definitions have fluctuated throughout the history of WS and have often been broken down into subcategories. In recent years, research generally applied the categories of DW, IW, and WW, however, with different names and subcategories, e.g., transport, travel, instruction, personal time, delay, etc. (Gong et al., 2011). In the current research, a six-category split is applied. One category of DW: (1) Production. Three categories of IW: (2) Transporting, (3) Preparing, and (4) Talking. Lastly, two categories of WW: (5) Walking and (6) Waiting.

As the WS methodology is based on direct observation of workers, data collection is costly, not scalable, and time-consuming (Zhao et al., 2019). A relatively large amount of observations are required to achieve statistical validity. According to Thompson (1987) a minimum sample size of 510 observations is required to achieve a 95% confidence level. If the aim is to analyze the distribution over time, the necessary number of observations per hour depends, among others, on the number of workers, and for 0-50 workers, 46 observations per hour are needed (CII, 2010). Data collection, therefore, often lasts from 3 to 5 days, where one or more observers must watch construction activities full time. Wandahl et al. (2022) concluded that if the observation requirements are obeyed, the sample can be robust and representative.

Because WS is time-consuming and costly, several researchers work on different approaches to automate the data collection by different kinds of vision-based and sensor-based technologies. Vision-based activity analysis requires single or multiple cameras for detecting and tracking resources as well as procedures for activity recognition (Liu and Golparvar-Fard, 2015). Sensor-based technologies enable the identification of measurement of workers' posture, motions, location, and presence (Cheng et al., 2017). Among the existing digital approaches for data collection, sensor-based technologies using body-worn sensors have gained greater attention among researchers for monitoring construction activities (Ryu et al., 2018) due to their flexibility to adapt to different external conditions and their reduced size easily to be embedded in, e.g., wristbands. Body-worn sensors can have integrated accelerometers, gyroscopes, and magnetometers, called Inertial Measurement Units (IMUs). IMUs can measure inertial body motions in three axes, as each activity creates unique acceleration signal patterns. Examples of such approaches could be smartwatches used for localization (Pérez et al., 2022) or for activity recognition or monitoring workers' health and safety conditions (Guo et al., 2017). Other more stationary sensors are location-based sensors like global navigations satellite system (Li et al., 2020), radio-frequency identification (Lu et al., 2011), ultrawideband (Teizer et al., 2020), and Bluetooth beacons (Görsch et al., 2022; Olivieri et al., 2017) which all can track workers' real-time location and automatically collect worker-activity-related data (Cheng et al., 2013).

The challenge of all these sensor-based approaches is that they do not directly substitute WS, as they do not provide the aforementioned share of time in different work categories. In particular, it is relevant for practitioners and researchers to know how much time is spent on DW. The lack of knowledge on a possible correlation between, e.g., presence, location, or worker movement and DW represents a gap in the current body of knowledge.

By adopting wrist-worn Global Navigation Satellite System (GNSS) sensors, it is possible to monitor workers' positions. Their position and time data can be used to identify how much they move in distance and duration. In the WS technique, these are related to categories 2 and 5, i.e., Transporting and Walking, respectively. This research is based on an underlying assumption that DW is negatively correlated to the workers' movement. This assumption seems logical, as the more time workers spend moving around on the construction site, the less time they can spend on value-adding activities. This assumption is based on the work by Neve et al. (2020b), who analyzed the correlation between all categories of WS in three different case studies, and concluded that the strongest correlation was between DW and Transporting + Walking combined. If this correlation is strong and valid, it would be possible to use sensor-based location data of how much construction workers move around as an indicator for their DW rate. Only when this correlation has been tested and confirmed does it makes sense to continue with the GNSS data collection approach for an automated WS data collection.

The aforementioned forms a gap in the current body of knowledge, and this research sets out to close this gap by answering the research question (RQ):

RQ: *What is the relationship between construction workers' Movement (Walking and Transporting) and time spent in Direct Work activities?*

The RQ will be answered and exemplified based on data from renovation projects. Renovation projects were selected for several reasons. Firstly, previous research (Neve et al., 2020b) has indicated that such a relationship exists in renovation projects partly because renovation projects have issues related to existing site conditions, which can create logistical challenges. Secondly, the logistical difficulties of renovation projects (e.g., Kemmer and Koskela, 2014) likely results in more worker movement. Lastly, several resources conclude that renovation often has lower productivity than other types of construction.

## 2. Methodology

The study adopted the Case Study method (Yin, 2003), as the primary research strategy because it enables an investigation of a given phenomenon. A case study is an empirical inquiry that investigates a contemporary phenomenon in depth and within its real-life context, especially when the boundaries between phenomenon and context may not be clearly evident. The phenomenon of the study comprises the relationship between construction workers' movement and time spent in DW. The real-life context is represented by five renovation projects located in five different cities in Denmark. The authors studied the phenomenon through the application of the WS technique.

The methodology section first introduces the five projects. Secondly, the data acquisition process is described. Thirdly, data aggregation and data cleaning are presented. Fourthly, it is explained how the linear regression data analysis was conducted. Lastly, the data discussion presents the assumption adopted in this study.

### 2.1. Description of Case Studies

Five cases of renovation projects were chosen, named in this research as Case 1, 2, 3, 4, and 5 (see Table 1). The cases were selected mainly based on three criteria: (1) they had to be renovation projects, (2) they needed to be similar to compare production system behaviors; and (3) the trades had to include traditional renovation work such as carpentry, painting, masonry, and so forth, which would occur on any renovation project.

**Table 1.** Case studies characterization.

| | Case 1 "Roskilde" | Case 2 "Odense" | Case 3 "Aarhus" | Case 4 "Vejle" | Case 5 "Herning" |
|---|---|---|---|---|---|
| Contract | General contr. | General contr. | Turnkey contr. | General contr. | General contr. |
| Value (mill $) | ~59 | ~73 | ~53 | ~59 | ~31 |
| Duration (years) | 5 | 5 | 4 | 4 | 5 |
| Built (Year) | 1950s | 1950s | 1960s | 1950s | 1950s |
| Size ($m^2$) | 48,000 | ~46,000 | 23,700 | 46,500 | 22,800 |
| Units (no) | 593 | 587 | 297 | 601 | 291 |

The chosen cases' original building structures and floor plans were very similar, and they were planned to go through comparable deep renovations, including the building envelope, interior, and installations. All cases were social housing renovation projects consisting of apartments (the details for each are outlined in Table 1). All cases were located in comparable cities in Denmark. A short description of each case is presented in the following.

*Case 1*: The case study was conducted in the city of Roskilde. The project consists of 24 five-story housing buildings. Four buildings were under renovation during the period of this case study. The main renovation tasks were carpenter work such as replacing windows, new facades, new roofs, etc. For this reason, most of the renovation activities were conducted outside the buildings from the façade scaffolding. Installing new ventilation and electrical systems were the only two indoor renovation activities. During the execution of the renovation project, tenants were granted rehousing in the period when their apartment was being renovated, but they were living in the apartment during the remaining renovation.

*Case 2*: The case study was conducted in the city of Odense. The project consists of two- to four-story buildings with two apartments on each floor. There is a total of 587 housing units. The buildings were first established in the early 1950s. The renovation included replacing old balconies, windows, kitchens, and bathroom interiors, adding insulation in walls, putting up drywall partitioning walls, and turning some units into accessible housing units by installing elevators in the stairwells. During the execution of the renovation project, tenants were rehoused in the period when their apartment was being renovated, but they were living in their apartment during the renovation of the neighboring buildings. To minimize the need for rehousing, only around 15% of the units were renovated simultaneously.

*Case 3*: The case study was conducted in the city of Aarhus. The project is a social housing complex initially built in the 1960s and now undergoing deep renovation, including new facades, new roof, and completely new installations. In total, there were 297 housing units. During the execution of the renovation project, tenants were rehoused in the period when their apartment was being renovated, but they were living in their apartment during the renovation of the neighboring buildings.

*Case 4*: The case study was conducted in the city of Vejle. The project is a 46,500 $m^2$ large social housing complex initially built in the 1950s and now undergoing a deep renovation, including new brick facades, new roof, new windows, additional insulation, new installations, and new interior like kitchen, bathroom, flooring, etc. During the execution of the renovation project, tenants were rehoused in the period when their apartment was being renovated, but they were living in their apartment during the renovation of the neighboring buildings.

*Case 5*: The case study was conducted in the city of Herning. The project consists of 350 housing units established between the years 1953-1957. In total, 19 blocks, all three stories high and with a basement. The building complex was undergoing a deep renovation where all units got a new kitchen, bathroom, facades, balcony, and completely new installations. All blocks got a new roof, improved insulation, and restored basements. Elevators were installed for 90 of the units, and several units were merged into larger units, resulting in 311 units after the renovation. During the execution of the renovation project, tenants were rehoused in the period when their apartment was being renovated, but they were living in their apartment during the renovation of the neighboring buildings.

### 2.2. Data Acquisition Description

This study adopted a six-work category classification during the WS technique. The categories were DW also called Production (Category 1). Three categories fall into IW, namely Transporting (Category 2), Preparing (Category 3), and Talking (Category 4). Finally, two categories of WW, namely Walking (Category 5) and Waiting (Category 6). Moreover, in this study, the time of each observation was recorded. For this, the authors adopted the smartphone application "Counter – Tally Counter" by Tevfik Yucek. This application allowed the researchers to digitally record each observation with an exact time stamp and export this data in a Comma-Separated Value (CSV) format for further processing.

For gathering data at the construction sites, the observers, i.e., the authors of this paper, conducted random tours under normal conditions, that is, representing standard workdays. Visual observations were conducted by observers walking the construction site from the beginning of the workday until the end of the workday (8 hours/each). The tours aimed to avoid observing patterns of behavior. Hence, the observers varied their routes through the job site and, to increase randomness, the times for observations. The data collection studies were conducted in different periods from 2017 to 2021 (see row 1 in Table 2). The duration of the studies was different from case to case. For each case, the necessary number of observations to achieve a 95% Confidence Interval was based on CII (2010). Hence, the total number of days of direct observations was 27 days, distributed as following (see row 2 in Table 2): 9 days in Case 1 and 2; 5 days in Case 3; and 3 days in Case 4 and 5.

**Table 2.** Information regarding Work Sampling.

| | Case 1 | Case 2 | Case 3 | Case 4 | Case 5 |
|---|---|---|---|---|---|
| Period of the study | Weeks 45 and 46, 2021 | Weeks 25 and 26, 2021 | Week 49, 2017 | Week 12, 2018 | Week 11, 2018 |
| Days of observation (days) | 9 | 9 | 3 | 3 | 3 |
| Number of workers observed (no) | 40 | 50< | 40> | 50< | 50< |
| Number of total observations (no) | 1,642 | 2,641 | 5,734 | 12,650 | 24,933 |
| Observations per hour (no) | ~23 | ~37 | ~44 | ~85 | ~83 |
| Min observations needed per hour (no) | 7.5 | 7.5 | 21 | 21 | 21 |
| Hourly datapoint (no) | 60 | 61 | 29 | 26 | 31 |
| Trades observed | Carpenter Mason HVAC Demolition Scaffolder | Carpenter Mason HVAC Electrician Demolition | Carpenter HVAC Painters Flooring Facade | Carpenter Mason HVAC Painters Concrete | Carpenter Mason Facade Flooring Demolition |
| Work Sampling data also used in | 10 | 8;9 | 1;3;4;5;6 | 1;3;4;5;6 | 1;2;3;4;5;6;7 |

1 (Neve and Wandahl 2018); 2 (Neve et al. 2020); 3 (Teizer et al. 2020); 4 (Neve et al. 2020); 5 (Wandahl et al. 2021); 6 (Neve et al. 2021); 7 (Johansen et al. 2021); 8 (Pérez et al. 2022); 9 (Salling et al. 2022); 10 (Pérez et al. 2022)

Five trades (see row 8 in Table 2) were observed during the WS application. The number of trades for each case was chosen to represent the majority of work in progress during the study periods, so the production system behavior of each case could be analyzed and compared. The number of workers observed during each case study varied but was always around 50 (see row 3 in Table 2). At the end of the job site visits, the number of observations in each case study was distributed as follows (see row 4 in Table 2): 1,642 observations in Case 1; 2,641 observations in Case 2; 5,734 observations in Case 3; 12,650 observations in Case 4; and 24,933 observations in Case 5. Cases 4 and 5 had a high number of observations, as several students were assigned as observers, e.g., in case 5, 12 students collected data together with the research team. The number of observations per hour (see row 5 in Table 2) was in the range of 23 to 85 observations per hour, and for all cases, this number was higher than the needed number of observations per hour (see row 6 in Table 2) to obtain a 95% Confidence Interval, calculated based on CII (2010).

To achieve uniform datasets among the five cases, all observations were grouped into hourly observations. The accumulative observations in each hour were then assigned as the dataset for further analysis in this research. Considering eight hours of daily observation, the number of hourly data points assigned to analyze further was N = 207 (see row 7 in Table 2), distributed as follows: 60 data points from Case 1= (representing 29% of the total sample); 61 from Case 2 (30%); 29 from Case 3 (14%); 26 from Case 4 (12%); and 31 from Case 5=31 (15%).

The results of the WS are shown in Table 3, including information on observation counts and percentages in each category and the 95% CI (plus-minus two times the standard deviation) for the DW category.

**Table 3.** Results of the WS for each case.

| Data | Prod. | Talk. | Prep. | Trans. | Walk | Wait | Total |
|---|---|---|---|---|---|---|---|
| Case 1 | | | | | | | |
| N | 411 | 136 | 419 | 320 | 230 | 126 | 1,642 |
| $\bar{p}$ | 25.03% | 8.28% | 25.52% | 19.49% | 14.01% | 7.67% | 100% |
| $\pm 2 \cdot s$ | ±2.2% | | | | | | |
| Case 2 | | | | | | | |
| N | 571 | 284 | 641 | 542 | 416 | 187 | 2,641 |
| $\bar{p}$ | 21.62% | 10.75% | 24.27% | 20.52% | 15.75% | 7.08% | 100% |
| $\pm 2 \cdot s$ | ±1.3% | | | | | | |
| Case 3 | | | | | | | |
| N | 1,982 | 733 | 1,439 | 833 | 510 | 237 | 5,734 |
| $\bar{p}$ | 34.57% | 12.78% | 25.10% | 14.53% | 8.89% | 4.13% | 100% |
| $\pm 2 \cdot s$ | ±1.3% | | | | | | |
| Case 4 | | | | | | | |
| N | 4,928 | 1,825 | 3,454 | 845 | 881 | 717 | 12,650 |
| $\bar{p}$ | 38.96% | 14.43% | 27.30% | 6.68% | 6.96% | 5.67% | 100% |
| $\pm 2 \cdot s$ | ±0.9% | | | | | | |
| Case 5 | | | | | | | |
| N | 7,777 | 6,259 | 4,684 | 2,314 | 1,896 | 2,003 | 24,933 |
| $\bar{p}$ | 31.19% | 25.10% | 18.79% | 9.28% | 7.60% | 8.03% | 100% |
| $\pm 2 \cdot s$ | ±0.6% | | | | | | |

The N value in Table 3 represents the number of observations in each work activity category. For example, of the 1,642 observations in Case 1 (see row 1 in Table 3), 411 were classified into the Production category, representing 25.03% of the total observations; 136 into the Talking category, representing 8.28%; 419 into the Preparing category, representing 25.52%; 320 into the Transporting category, representing 19.49%; 230 into the Walking category, representing 14.01%; and the remaining 126 into the Waiting category, representing 7.67%.

### 2.3. Data Aggregation and Data Validation

During the data aggregation process, the observations were grouped into hourly intervals, which then are the data points used in the analysis. For the data cleaning, only data covering an entire hour were used. An example is on the third day of data collection in Case 5, where work stopped at 14:30. In this situation, all observations from 14:00 to 14:30 were deleted as they did not cover an entire hour.

Table 4 shows the average numbers of all six data sets, the five case studies, and the sum of all data points, named in this study as Case ALL, resulting in N = 207. The complete data set of the aggregated data is presented in Appendix 1.

**Table 4.** Average aggregated data points from each case.

| Data sets | Prod. avg. | Talk. avg. | Prep. avg. | Trans. avg. | Walk avg. | Wait avg. |
|---|---|---|---|---|---|---|
| Case 1 (N=60) | 26.07% | 8.66% | 25.54% | 16.97% | 14.39% | 6.70% |
| Case 2 (N=61) | 21.50% | 10.94% | 24.39% | 20.05% | 15.93% | 7.18% |
| Case 3 (N=29) | 36.65% | 15.15% | 25.30% | 6.62% | 11.50% | 4.78% |
| Case 4 (N=26) | 35.83% | 15.08% | 24.53% | 12.81% | 8.45% | 3.30% |
| Case 5 (N=31) | 28.80% | 23.84% | 21.20% | 8.50% | 9.33% | 8.33% |
| Case ALL (N=207) | 29.77% | 14.73% | 24.19% | 12.99% | 11.92% | 6.06% |

After the aggregation and the cleaning, each data point is considered an array that includes the relative percentage of observation in each category for that hour. For example, in Case 1, datapoint no 14 is the observations on day 2 from the time 13.00-14.00, and the value of the datapoint is [Production = 35.29%; Talking = 17.65%; Preparing = 5.88%; Transporting = 17.65%; Walking = 5.88%; Waiting = 17.65%]. This adds up to 100% and covers 17 observations made in that time interval. The average of these data points is illustrated in Table 4.

In addition, stabilization curves of the share of observations of the DW were created to provide a visual check of the accuracy of the collected data, cf. Fig. 1. All WS data are assessed as valid based on (a) the stabilization curves (Fig. 1) being stable after around 50% of the data collection; (b) more data points than the minimum required (CII, 2010) were collected (Table 2); (c) the calculated 95% confidence interval is low, i.e., around 1 percent point (Table 3).

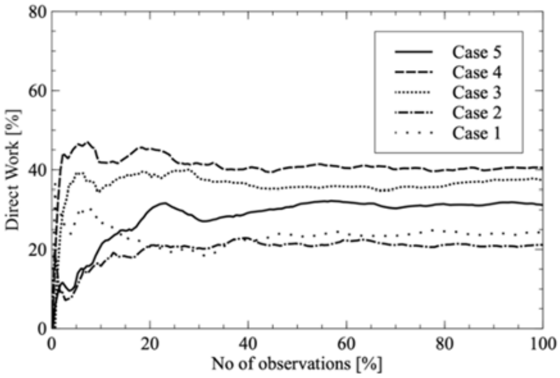


**Fig. 1.** Stabilization curves of the five cases.

### 2.4. Data Analysis

The data analysis aimed to test whether a possible relationship between workers' movement and DW is statistically significant. DW is the Production category, and Movement is the sum of Walking and Transporting in the WS data. The analysis was applied to the six data sets representing Case 1-5 and Case ALL. The authors used the Statistical Package for the Social Sciences (SPSS) software for statistical analysis. The authors selected a combination of four quantitative techniques in order to answer the research question: (1) curve estimation; (2) linear regression; (3) ANOVA analysis; and (4) t-test.

1. *Curve estimation* with 11 equations was applied, first to understand if a statistically significant relationship can be established for the six different datasets, and second to understand which equation provides the best predictive capabilities. The 11 equations were linear, logarithmic, inverse, quadratic, cubic, compound, power, S, growth, exponential, and logistic. The linear was found best fitting, thus the introduction of the last three tests.

2. *Linear regression analysis* providing a linear equation.

3. *ANOVA analysis* providing a p-value that reveals the statistical significance of the linear regression model's predictive capabilities.

4. *t-test* enabling the calculation of 95% Confidence Intervals (CI) for the linear regression model's coefficients.

Common for the two first tests are that they rely on interpreting the correlation coefficient (R). Previous recommendations (Cohen, 1988) outline that R > 0.5 reflects a large effect size. Research in the same area as this has previously used R = 0.318 as an acceptable level (Liu et al., 2011). Nonetheless, in this research R = 0.5 is chosen as the minimum limit for accepting any relationship established through the statistical analysis. The R-value can be squared ($R^2$) to instead reflect the predictive capabilities of the independent variable in the analysis. The $R^2$ value corresponding to R = 0.5 is 0.25, and thus, $R^2 = 0.25$ is the lower acceptance limit. Additionally, all established relationships must have a statistical significance level above 95% ($p \leq 0.05$) to be valid.

### 2.5. Data Discussion

The last section of this paper discusses the correlation of the result considering three assumptions described in the following.

*1. Structure of the day.*

Like most previous WS research, this research assumes that a working day is homogenous, that is, the hourly interval in the results is identical in terms of work, working conditions, etc. In general, WS results provide a distribution of work time in different categories, cf. section 2.2, assuming that this distribution is valid at all times for the phenomenon observed. A few previous WS studies have indicated that this is not a fully valid assumption. Neve et al. (2020b) presented detailed day curves, which showed a fluctuation of DW during the day. In addition, Björkman et al. (2010) and Gouett et al. (2011) concluded that especially starts and stops during the day influenced the DW.

*2. Movement vs. transporting and walking.*

The correlation in the results is analyzed with the assumption that DW is correlated with Transporting and Walking combined, called Movement. This assumption was based on the work by Neve et al. (2020b), which analyzed the correlation between all categories of WS in three different case studies. They concluded that the strongest correlation was between DW and Transporting + Walking combined. However, that assumption has not been tested.

*3. Global vs. local.*

An underlying assumption in the research question is that the correlation between DW and Movement is global, i.e., valid across cases. Wandahl et al. (2021) investigated a sample of 474 WS studies from the last 50 years and concluded that academics and practitioners should be careful when using WS for generalizing purposes. Josephson and Björkman (2013) concluded that WS was not valid to generalize over time, as the working conditions can change and will influence the WS result.

In addition to discussing the above-mentioned three assumptions, the discussion will also consider the topics of correlation vs. causality in the research, limitation of the research, and industrial implication of the research.

## 3. Results

This section presents the results of the four quantitative techniques chosen to answer the research question.

### 3.1. Curve estimation

To reach this research's objective, curve estimation was done on the six data sets (cf. Table 5).

**Table 5.** Curve estimations of the six data sets with 11 equations.

| Data sets | | Linear | Logar. | Inverse | Quadr. | Cubic | Comp. | Power | S | Growth | Expon. | Logistic |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Case 1 (N = 60) | $R^2$ | .119 | .0052 | .014 | .122 | .132 | .004 | .072 | .101 | .004 | .004 | .004 |
| | p | .007 | .08 | .0364 | .025 | .047 | .619 | .038 | .014 | .619 | .619 | .619 |
| Case 2 (N = 61) | $R^2$ | .451 | .553 | .185 | .519 | .582 | .395 | .368 | .109 | .395 | .395 | .395 |
| | p | .001 | .001 | .001 | .001 | .001 | .001 | .001 | .009 | .001 | .001 | .001 |
| Case 3 (N = 29) | $R^2$ | .356 | .331 | .262 | .368 | .398 | .433 | .229 | .148 | .433 | .433 | .433 |
| | p | .001 | .001 | .005 | .003 | .005 | .001 | .009 | .039 | .001 | .001 | .001 |
| Case 4 (N = 26) | $R^2$ | .310 | .283 | .169 | .337 | .453 | .206 | .199 | .128 | .206 | .206 | .206 |
| | p | .003 | .005 | .037 | .009 | .004 | .020 | .022 | .072 | .020 | .020 | .020 |
| Case 5 (N = 31) | $R^2$ | .332 | .311 | .222 | .334 | .336 | .319 | .280 | .185 | .319 | .319 | .319 |
| | p | .001 | .001 | .007 | .003 | .010 | .001 | .002 | .016 | .001 | .001 | .001 |
| Case ALL | $R^2$ | .328 | .288 | .008 | .352 | .362 | .006 | .010 | .009 | .060 | .060 | .060 |
| (N = 207) | p | .001 | .001 | .001 | .001 | .001 | .001 | .150 | .171 | .001 | .001 | .001 |

The curve estimation results in the predictive capability ($R^2$) with the lower limit at $R^2 = 0.25$ and a statistical significance level at 95% ($p \leqq 0.05$) of each established relationship for the 11 equations. Table 5 reveals that only Case 1 fails in establishing a sufficient predictive capability with adequate significance. It can, therefore, be concluded that, a statistically significant relationship exists between DW and Movement for most of the 11 equations.

In Case 2 and Case 5, only the S and the Inverse function fail to reach the $R^2$ and the p threshold. For Case 3, only the S and the Power equations fail to reach the threshold. In case 4 and the Case ALL, only four of the 11 equations succeed in establishing a relationship above $R^2 = 0.25$ with a statistical significance level of 95% ($p \leqq 0.05$). The four equations are the linear, logarithmic, quadratic, and cubic. Thus, these four equations need closer analysis to conclude which is the best to explain the relationship between DW (%) and Movement (%).

### 3.2. Linear regression

Looking at the $R^2$ value for the six cases per the four selected equations, the Cubic is the best ($R^2$=0.426), followed by Quadratic ($R^2$=0.382), and then Linear ($R^2$=0.355). The ranking of the significance level is the opposite, where the Linear has the lowest average 95% significance value, and the Cubic has the highest.

The selected four equations' capacities are evaluated regarding both overfitting and underfitting, which is a known approach from machine learning. An overfit model is one that is too complicated for the data set. The regression model then becomes tailored to fit the quirks and random noise of the data set. Underfitting is the opposite. A visual inspection often reveals overfitting. Therefore, the four regression models and the data set are plotted in Fig. 2.

When DW increases to above 50%, the logarithmic model predicts that Movement is almost steady around 20%. This is clearly a false prediction, which is clearly seen in Fig. 2, and also logical in a causal view, as Movement must approach 0% when DW approaches 100%. The quadratic regression model predicts that a horizontal tangent line occurs, and Movement increases when DW increases. This is neither causal nor visible in the dataset in Fig. 2. Based on Fig. 2, it, therefore, becomes clear that both the logarithmic and the quadratic regression models are overfitting. The cubic regression model predicts a rather complicated relationship between DW and Movement, which is not easily explained with rational thinking or through a causal relationship. Based on the above argumentation, the linear regression model has the best and most valid predictive capability.

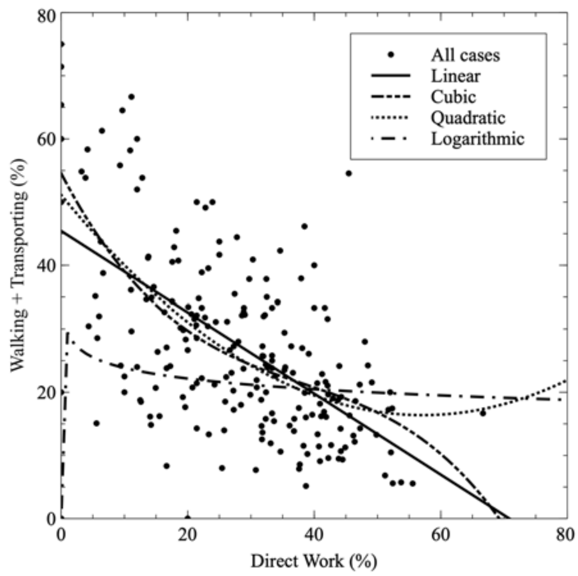


**Fig. 2.** Regression with four selected equations (N = 207).

### 3.3. ANOVA analysis

ANOVA is carried out to further investigate the linear regression model with DW as the independent (predictor) variable and Movement as the dependent (response) variable.

**Table 6.** Regression analysis, t-test (95% CI), and ANOVA for DW-Movement.

| Case | Regression y=ax+b | a (95% CI) | b (95% CI) | $R^2$ | ANOVA p-value |
|---|---|---|---|---|---|
| Case 1 N=60 | y = -0.407x + 0.420 | (-.698; -.115) | (.333; .506) | .119 | .007 |
| Case 2 N=61 | y = -0.848x + 0.542 | (-1.093; -.604) | (.482; .603) | .451 | .001 |
| Case 3 N=29 | y = -0.526x + 0.374 | (-.805; -.246) | (.265; .483) | .356 | .001 |
| Case 4 N=26 | y = -0.344x + 0.336 | (-.560; -.128) | (.256; .416) | .310 | .003 |
| Case 5 N=31 | y = -0.369x + 0.285 | (-.568; -.170) | (.223; .346) | .332 | .001 |
| ALL N=207 | y = -0.641x + 0.454 | (-.767; -.515) | (.415; .494) | .328 | .001 |

Table 6 presents the linear regression analysis for the five cases, including the accumulative case (named ALL).

A t-test gave the 95% CI for the predictor coefficient (a) and the constant coefficient (b), and the $R^2$ values from the regression analysis, and finally, the ANOVA result shows the statistical significance level for each case.

Further, Fig. 3 visually represents the linear regression models with the data from all five cases. The bold line in Fig. 3 illustrates the summarized regression model for all cases, whereas the thin lines represent the regression model for each case. The predictive capability of the linear model range depends on the case and ranges from $R^2=0.119$ to $R^2=0.451$, with an average of $R^2=0.328$. The significance level based on the ANOVA process shows p-values ranging from $p=0.001$ to $p=0.007$. Thus, all are statistically significant above the 95% level.

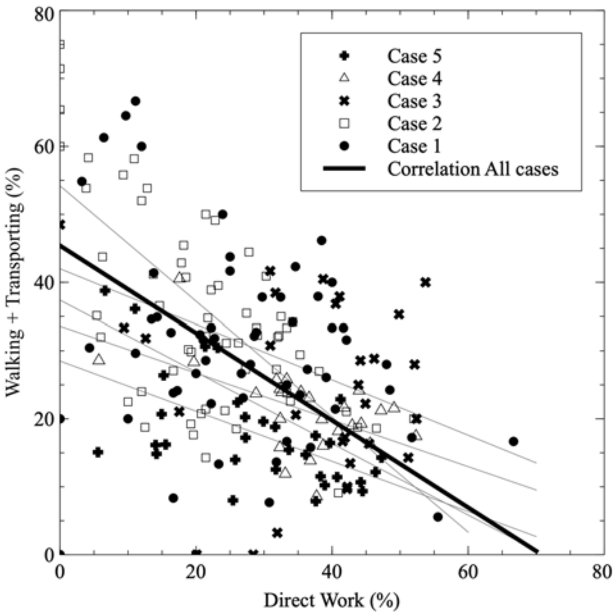


**Fig. 3.** Linear regression per case (N = 207).

### 3.4. T-test Analysis

Based on the t-test, the 95% CI for parameters a and b in the linear regression model is calculated. The result is shown in Table 6. The linear regression model equation for the accumulative Case ALL is presented in equation 1.

$$y = -0.641x + 0.454 \tag{1}$$

*where:*
*y = Movement (%)*
*x = DW (%)*

It is not expected that the regression model would be able to predict all the new data points, as R only is R=57.27%.

## 4. Discussion

Several aspects of the observed correlation can be discussed to shed further light on how DW and Movement of construction workers are connected.

### 4.1. Structure of A Day

The correlation in the results is analyzed with the assumption that a working day is homogenous, i.e., each hourly interval in the analysis is identical in terms of work, working conditions, etc. Previously WS has clearly demonstrated that this is not a fully valid assumption. Neve et al. (2020b) present detailed day curves, which show, cf. Fig. 4, a clear pattern of the amount of DW fluctuating during the day.

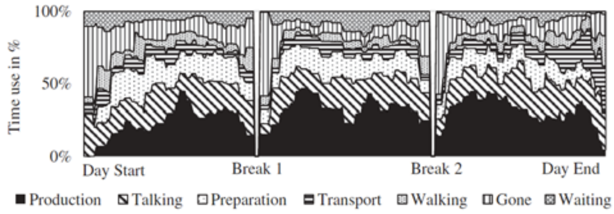


**Fig. 4.** Day curve example (Neve et al., 2020b).

The fluctuation is mainly due to productivity issues at the start of the day, just before and after breaks, and at the end of the day (Johansen et al., 2021; Neve et al., 2020b). It seems that the DW-Movement ratio is relevant to investigate further with the present data with respect to time of the day. Fig. 5 depicts both the above-mentioned ratio and the $R^2$ value for each hourly interval for Case 1-5.

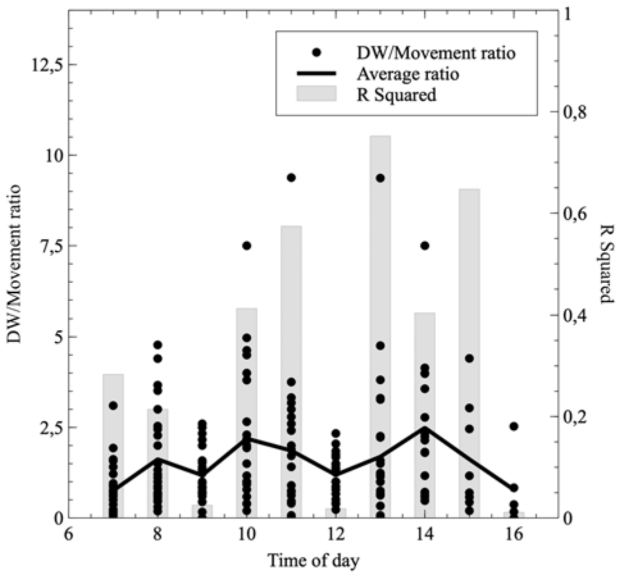


**Fig. 5.** Structure of a day (N = 207).

Fig. 5 reveals some very interesting insights. The dots on the figure represent the ratio of DW and Movement for each hourly interval (N = 207) of the five cases. In all five cases, there was a morning break around 9 o'clock and a lunch break around 12 o'clock. The bold line is the average ratio within each hourly interval. Firstly, the plot in Fig. 6 clearly shows the time of the breaks. These are the times where the DW-Movement ratio is lowest, i.e., low amount of DW and more Movement. The same issue is for the start of the day, where less DW is done, and more walking is needed to start the production, e.g., getting material, tools, walking to office trailers, etc. Also, at the end of the day, less DW and more Movement take place when shutting down the production. It is observed that this pattern is recognizable in the $R^2$ values. Around breaks, in the morning and the afternoon, the $R^2$ is very low, concluding that there is a weak correlation between DW and Movement. It is also clear that the $R^2$ is very high when production is running, i.e., in between breaks. In these periods, $R^2$ is higher than 0.5, showing that the more than 70% of the data can be predicted. This is a very strong correlation. It can, thus, be argued that the correlation between DW and Movement is very strong when normal production conditions exist.

### 4.2. Movement vs. Transporting and Walking

The correlation in the results is analyzed with the assumption that DW is correlated with Movement. This assumption is based on the work by Neve et al. (2020b), who analyzed the correlation between all categories of WS in three different case studies, and concluded that the strongest correlation was between DW and Transporting + Walking combined.

Fig. 6 depicts on the left side a linear regression model of DW and Walking and on the right side a linear regression model of DW and Transporting. The Case ALL data set is used, thus, N = 207 in both models, cf. Table 4 and Appendix 1.

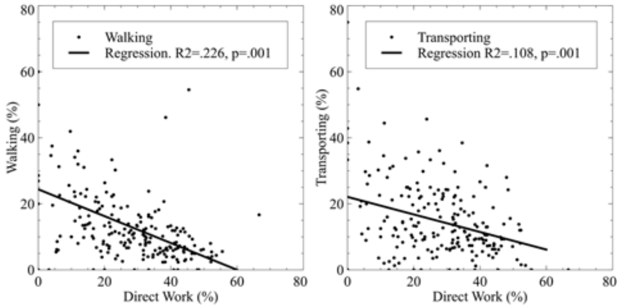


**Fig. 6.** Walking (left) vs. Transporting (right) (N = 207).

It is evident from the regression analysis depicted in Fig. 6 that both correlations are weak, i.e., $R^2 = 0.226$ and 0.108, respectively. These are weaker correlations than the Walking and Transporting combined (called Movement), cf. Table 6, where $R^2 = 0.328$. To further understand why the correlation is stronger when Walking and Transporting are combined than individually, one must understand what Walking and Transporting are on the construction site.

In WS, Walking describes a worker walking to or from the production area without tools or material in his hands. There are several typical causes for walking. It can be related to (a) getting to and from the work area in the morning, around breaks, or in the afternoon; (b) personal time for toilet breaks; (c) request for information, where a worker decides to walk to ask the foreman, site manager; (d) logistics, i.e., walking to pick up material, other supplies, and equipment and tools; (e) Unnecessary and extra breaks, where the worker walks to a site container, their car; (f) Unnecessary and extra walking in relation to c and d due to poor planning and management; and (g) Apparent efficiency where workers walk around to deceptively show that they are busy and efficient.

Transporting is when construction workers are observed walking with material, supplies, or tools in their hands. Transporting is, therefore, very related to points d and f above. Thus, a dual relationship is often the case. That is, when a worker is walking due to logistic tasks, the worker will soon thereafter be doing transporting, i.e., walking back to the work area with material, supplies, or tools. Thus, a causal explanation for why it makes sense to combine Walking and Transporting into Movement is present.

### 4.3. Global vs. Local Correlation

This research assumed that the correlation between DW and Movement was global, i.e., valid across cases. Table 5 showed that the correlation was significant in four of the five cases. Table 6 revealed that the a and b parameters in the linear correlation varies from case to case. It seems, therefore, that one cannot assume a certain correlation globally. Instead, the correlation is present only locally.

Table 7 shows the global predictability of each of the cases on each other. The data show the percentages of the data points that are inside the 95% CI of the predictor regression model. For example, the regression model of Case 2 (y = -0.848x + 0.542) can predict 58% of the data points of Case 5.

**Table 7.** Predictive capability globally.

| | Case 1 | Case 2 | Case 3 | Case 4 | Case 5 | *Average predictive capability* |
|---|---|---|---|---|---|---|
| Case 1 | 100% | 39% | 14% | 35% | 16% | *26%* |
| Case 2 | 53% | 100% | 76% | 73% | 58% | *65%* |
| Case 3 | 40% | 26% | 100% | 62% | 65% | *48%* |
| Case 4 | 25% | 20% | 34% | 100% | 29% | *27%* |
| Case 5 | 18% | 15% | 69% | 19% | 100% | *30%* |
| *Average prediction of case* | *34%* | *25%* | *48%* | *47%* | *42%* | |

Table 7 shows that the predictive capability of each linear regression model ranges from 14% to 76% of data points. On average, the predictive capability is a bit lower than 50%. The predictive capability of each case's linear regression model fluctuates; thus, no conclusion can be drawn regarding the global correlation.

Interestingly, each case is somewhat different in how good it is at being predicted, cf. the bottom row in Table 7. It seems that Case 2 is harder to predict than, e.g., Case 3. Therefore, the homogeneousness of Case 2 and Case 3 is investigated. From Table 1, it is seen that the two cases are similar in trades, type of project, and year build but quite different in size, i.e., the number of units to be renovated. The context of the two cases is thus concluded to be homogenous. The performance data of the two cases are reviewed in Fig. 7.

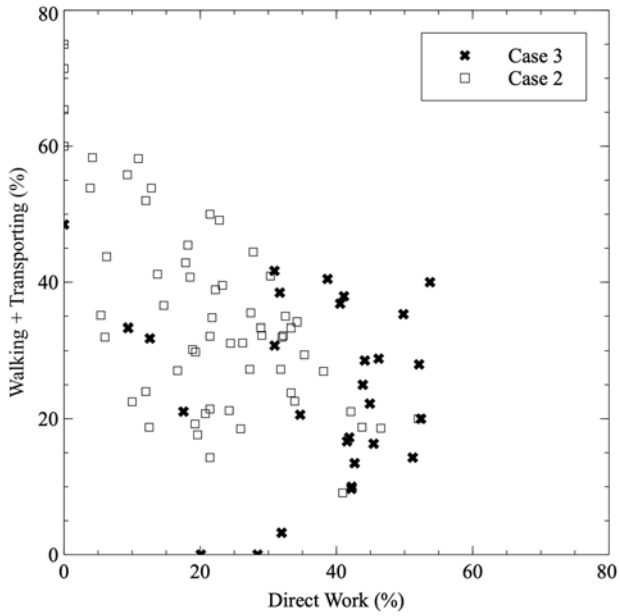


**Fig. 7.** Case 2 and 3 datapoints comparison.

From Fig. 7, it is seen that there is a significant difference in performance, i.e., DW, of Case 2 and Case 3, where Case 3 has DW rates of around 40%, and Case 2 has DW rates in the range of 20%. This could indicate that the predictive capability is not equally good on the entire range of DW. This is investigated further. The regression model of Case 1 (cf. Table 6) is used to predict the 61 data points of Case 2, the 29 data points of Case 3, the 26 data points of Case 4, and the 31 data points of Case 5 (cf. Table 4). Then, the regression model of Case 2 is used to predict the data points of Case 1 to 5, and so on. This results in N = 828 (4*207=828) predictions of Movement (y in the equation in Table 6) based on DW (x in the equation in Table 6). All the N = 828 predictions are attached in Appendix 2.

A prediction is accepted if the result of the linear regression equation is within the range of Movement plus minus the CI of the actual datapoint. The result of this analysis is depicted in Fig. 8, where the average predictability relative to the DW is illustrated.

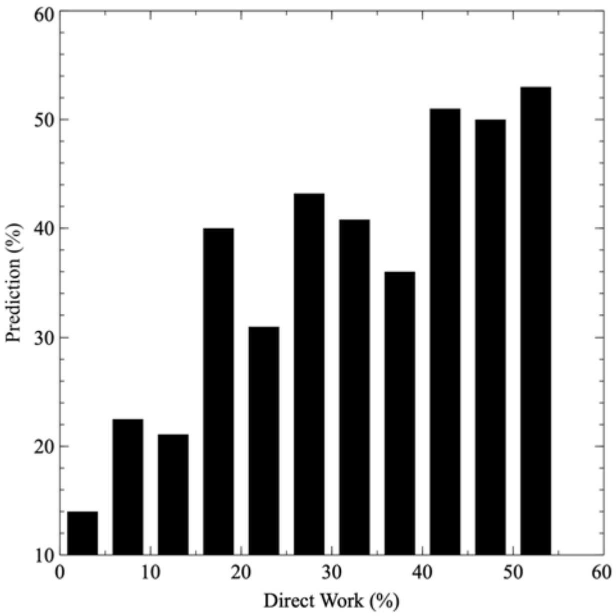


**Fig. 8.** Predictability within DW intervals (N = 828).

There is a clear trend visible in Fig. 8. The lower DW is, the harder it is to predict Movement. When DW is low, there are two possible explanations: (1) if the production activities are dependent on handling (e.g., carpentry work), it indicates that the onsite production is running poorly; and (2) if the production activities are industrialized (e.g., off-site fabrication), it reduces the labour workload and, consequently, it increases the time spent on transporting and handling the pre-fabricated systems. Another possible reason can be on-site logistics in terms of Movement, but a range of other factors could also be the cause for low DW. For instance, lack of materials, incomplete drawings, work interference, out-of-sequence work, etc., are often mentioned as disablers of high productivity (Hughes and Thorpe, 2014; Jarkas, 2015). It is, therefore, reasonable that the model's predictability is low on the lower end of the DW scale. Fig. 8 illustrates that the predictability is around 30-50% as long as DW is above 15%. This supports the conclusion that it is globally accepted that Movement correlates with DW in non-industrialized systems, however, the linear regression equation with parameters a and b is only valid locally.

### 4.4. Correlation or Causality

This research uses a statistical approach to analyzing WS data from five cases. The main statistical method is, therefore, the regression model, which is looking for the correlation between two variables. In this research, the variables are DW as the independent predictor variable and Movement as the dependent response variable. Examples of statistical analysis of WS data can be found in, e.g., Siriwardana et al. (2017) and Jenkins and Orth (2004), who both conducted regression analysis between DW and productivity. Neve et al. (2020b) performed statistical regression on the different WS categories on single projects, and Wandahl et al. (2022) also applied statistical measures to a single WS study. Especially the statistical correlation between DW and Construction Labor Productivity has been heavily debated in academia (Gouett et al., 2011; Shahtaheri et al., 2015). The debate centers around whether there is a causal relationship between the time you spend on productive activities and the quantitative output of the productive activity. The conclusion in the DW vs. Productivity debate is, as in all other statistical cases, that the researchers need to confirm that a causal relationship between the two variables exists in order to draw a conclusion with practical implications. The following is, thus, a discussion of whether a causal relationship between DW and Movement exists.

Section 4.2 explained how Movement was composed of Walking and Transporting combined, what Walking and Transporting are on a construction site, and how there is a logical dependency between Walking and Transporting. Movement is thus the share of time construction workers use on walking empty-handed or with material or tools. DW, on the other hand, is the share of time construction workers uses in production. In WS, time is considered as either DW, IW (e.g., Transporting), or WW (e.g., Walking), and the sum of these three ads up to 100%. It is thus evident that if more time is spent on DW, less time is spent on IW and WW, illustrating a logical negative correlation. Moreover, the average time spent on Movement is 27.6% of the work time (cf. Appendix 1), thus the largest none productive category.

When production is planned well and execution is running smoothly with good flow and high efficiency, it is logical that a worker not doing value-adding activities is likely to do walking or transporting material or tools. Some walking is probably due to other activities like personal time, going for breaks, etc. In this high efficiency scenario, a sound negative correlation between DW and Movement is causal. This is also visible in Fig. 8, where predictions of DW based on Movement are best when DW is high. When production is running poorly, there are likely more causes for Movement, like rework, additional Movement due to on-site logistics, more waiting time, where workers can walk around, etc. Again, this reasoning is supported by Fig. 8, where it is evident that the correlation between DW and Movement is lowest when DW is low.

The above cause and effect discussion needs to consider that construction production here is viewed as a socio-technical system. This implies firstly that, in system thinking, it is likely that not modeled or not identified components exist. In other words, there are probably several other reasons for Movement than not producing. This research has limited its scope to not search for and include these potential other factors causing Movement. Therefore, achieving a significant $R^2$ effect size cannot be expected, which is also the case for this research, where the overall correlation between DW and Movement has an effect size of $R^2 = 0.328$.

### 4.5. Limitations

This research holds several limitations. Firstly, all of the collected data originates from Danish renovation projects. Renovation projects distinguish themselves from other project types by having several unique characteristics. Recent works by Neve and Wandahl (2018), and Tzortzopoulos et al. (2020), have shed light on renovation and point out that the main challenges are: (1) existing building structures with several unknown characteristics; (2) an often not optimal construction site layout for logistics and material handling; (3) highly specialized tasks and trades, e.g., removing asbestos, etc.; and (4) dealing with occupied buildings and tenants on site. Tenants in proximity to ongoing construction work require a high level of protection (e.g., dust, noise) and make managing the craftsmen-tenant relationship an additional challenge in renovation projects. These combined characteristics make renovation a more challenging environment to manage than, e.g., new build. Thus, transferability to other types of construction is limited.

Secondly, the data analysis only includes five case studies. To generalize based on five cases contains uncertainties, which should be considered when interpreting the conclusion of this research. The issue of global correlation has been discussed in this research, and it is clear that the five cases behave differently, as all cases show a correlation but with different regression models. However, it is still possible to conclude positively on the research question targeted in this research. Moreover, to include more cases would likely not change the conclusion, as these potential additional cases would also be different but show a correlation between DW rates and Movement.

Thirdly, the effect size R, describing the strength of the relationship between the two variables DW and Movement, varies from 0.34 to 0.67 on a scale from 0 to 1, where 0 is no relationship at all between the two variables and 1 is a perfect correlation. The average effect size for all data is R = 0.57. R values were calculated from $R^2$ values in Table 6. It is often discussed in the literature what an acceptable effect size is, and according to (Cohen, 1988), an effect size of $R > 0.5$ is a large effect size. This view is also supported by Neve et al. (2020a), where a correlation with an effect size of R=0.49 is concluded as acceptable. When interpreting the effect size of the correlation between two variables, one needs to bear in mind that the two variables, in this case, originate from a socio-technical system. A socio-technical system does not behave in accordance with the law of physics as several parts of the system can influence the variables, and also, more important, human behavior impacts the system in a somewhat unpredictable manner. Thus, an effect size of R = 1 would never be possible in a socio-technical system (Wandahl et al., 2022). As a result, this research holds the limitation that it will not be possible to achieve a 100% correlation between DW and Movement thus, using Movement as an indicator for DW rates will inherently include some uncertainties.

### 4.6. Industrial Impact

The construction industry continuously strives to reach higher productivity rates to secure more value for money for the clients and the society and better profit margins for the supply team. One Lean approach to reach this is to be more effective, i.e., have less waste and more VAW time. WS is an excellent diagnostic tool for identifying how time is spent on the construction site, and the output of WS is a benchmark of how time is distributed. This benchmark is needed to identify the right countermeasures for reducing waste time. WS has, however, a downside, namely the effort and time required to do observations. Today, a contractor typically sets aside one site manager for five days to collect data. This large effort is often a barrier to conducting WS, which is why several different studies currently attempt to automate the data collection by means of different sensor-based technologies. The hope is that if technology can automate the WS data collection, more contractors would do it, and then likely more contractors would succeed in reducing waste time and then eventually improve construction labor productivity to the benefit of construction in general.

## 5. Conclusion

This research addressed the RQ of what is the relationship between construction workers' Movement (Walking and Transporting combined) and time spent in DW activities. The authors presented different possible analyses using the data gathered through the application of the WS technique in five case studies.

The authors selected a combination of four quantitative techniques in order to answer the research question, those being: (1) curve estimation; (2) linear regression; (3) ANOVA analysis; and (4) t-test. Then, this paper discusses the correlation of the result considering three assumptions: (1) the structure of the day; (2) global vs. local; and (3) Movement vs. Transporting and Walking.

The results of curve estimation were that four curves (linear, logarithmic, quadratic, and cubic) had $R^2$ above the threshold limit of $R^2 = 0.25$ and a statistically significantly better than $p \leqq 0.05$. The more detailed regression analysis showed that the logarithmic and the quadratic regression models are overfitting, and the cubic regression model predicts a complicated and unrealistic relationship between DW and Movement. Therefore, the linear regression model was selected for the further analysis. The ANOVA analysis showed that the predictability of the linear regression model on the different cases was $R^2 = 0.328$ with $p \leqq 0.001$. The t-test, was used to calculate the 95% CI for parameters a and b in the linear regression model.

This research raised topics to be examined in greater depth in future research efforts. This research work discusses the correlation of DW and Movement based on some assumptions previously mentioned. Future studies should consider other assumptions, such as (1) considering the structure of the day dissimilar regarding the features of the work activities conducted in each period of the day; for that, future studies should compare the WS results between cases exclusively among the same hourly interval; and (2) assuming the correlation between DW and Movement was local, in other words, that the share of time in different categories varies from each construction site to each construction site with different features.

Another topic for future investigation is to explore the correlation of DW and Movement in industrialized construction projects. This research aimed to understand the correlation of those two variables based exclusively on renovation projects where most of the activities conducted are extremely labour resources dependent. Further research can explore the variables collected in other construction projects, which present off-site construction. Examples of this type of project could be projects that adopt prefabrication, panelization, or modularization systems. Those systems will deliver the produced off-site components, consequently affecting the productivity of the whole project by reducing the share of time spent on onsite production activities.

## 6. Acknowledgement

This work was supported by the Independent Research Fund Denmark (grant no. 0217-00020B). The authors are also grateful for all the construction companies that opened their construction sites for data collection.

## 7. Institutional Review Board Statement

This research has been approved by the Aarhus University Institutional Review Board on the 11th March 2021 with approval ID 2020-09.

## 8. References

Barnes, R. M. (1968). *Motion and Time Study: Design and measurement of work.* Wiley.

Björkman, L., Josephson, P. E., and Kling, R. (2010). *Arbetstidens använding vid VVS-montage - en fråga om struktur och ledarskab*. SBUF, Sweden.

Cheng, C. F., Rashidi, A., Davenport, M. A., and Anderson, D. V. (2017). Activity analysis of construction equipment using audio signals and support vector machines. *Automation in Construction*, 81, 240-253. doi:10.1016/j.autcon.2017.06.005

Cheng, T., Teizer, J., Migliaccio, G. C., and Gatti, U. C. (2013). Automated task-level activity analysis through fusion of real time location sensors and worker's thoracic posture data. *Automation in Construction*, 29, 24-39. doi: 10.1016/j.autcon.2012.08.003

Choy, E. and Ruwanpura, J. Y. (2006). Predicting construction productivity using situation-based simulation models. *Canadian Journal of Civil Eng.*, *33*(12), 1585-1600. doi: 10.1139/l06-088

CII. (2010). *A Guide to Activity Analysis*. University of Texas.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates.

Gong, J., Borcherding, J. D., and Caldas, C. H. (2011). Effectiveness of craft time utilization in construction projects. *Construction Management and Economics*, 29 (July), 737-751. doi: 10.1080/01446193.2011.595013

Gouett, M. C., Hass, C. T., Goodrum, P. M., and Caldas, C. H. (2011). Activity analysis for direct-work rate improvement in construction. *Journal of Construction Engineering and Management*, 137(12), 1117-1124. doi: 10.1061/(ASCE)CO.1943-7862.0000375

Guo, H., Yu, Y., Xiang, T., Li, H., and Zhang, D. (2017). The availability of wearable-device-based physical data for the measurement of construction workers' psychological status on site: From the perspective of safety management. *Automation in Construction*, 82, 207-217. doi: 10.1016/j.autcon.2017.06.001

Görsch, C., Al Barazi, A., Seppänen, O., and Abou Ibrahim, H. (2022). Uncovering and Visualizing Work Process Interruptions Through Quantitative Workflow Analysis. *Proceedings of IGLC30*, Edmonton, Canada.

Hughes, R. and Thorpe, D. (2014). A review of enabling factors in construction industry productivity in an Australian environment. *Construction Innovation*, 14(2), 210-228. doi: 10.1108/CI-03-2013-0016

Jarkas, A. M. (2015). Factors influencing labour productivity in Bahrain's construction industry. *International Journal of Construction Management*, 15(1), 94-108. doi: 10.1080/15623599.2015.1012143

Jenkins, J. L. and Orth, D. L. (2004). Productivity improvement through work sampling. *Cost Engineering*, 46(3), 27-32.

Johansen, P., Christensen, S., Neve, H. H., and Wandahl, S. (2021). Lean renovation – a case study of productivity, flow, and time improvements. *Proceedings og IGLC29*, Lima, Peru.

Josephson, P. and Björkman, L. (2013). Why do work sampling studies in construction? The case of plumbing work in Scandinavia. *Engineering, Construction and Architectural Management*, 20(6), 589-603. doi: 10.1108/ECAM-12-2011-0108

Kalsaas, B. T. (2011). On the discourse of measuring work flow efficiency in construction. A detailed work sampling method. *Proceedings og IGLC29*, Lima, Peru.

Kemmer, S. and Koskela, L. (2014). Understanding production management of refurbishment projects of a housing association – an exploratory case study. *Proceedings of IGLC22*, Oslo, Norway.

Li, C. T., Cheng, J. C. P., and Chen, K. (2020). Top 10 technologies for indoor positioning on construction sites. *Automation in construction*, 118, 103309. doi: 10.1016/j.autcon.2020.103309

Liu, A. M. M., Ballard, G., and Ibbs, W. (2011). Work flow variation and labor productivity: Case study. *Journal of Management in Engineering*, 27(4), 236-242. doi: 10.1061/(ASCE)ME.1943-5479.0000056

Liu, K. J. and Golparvar-Fard, M. (2015). Crowdsourcing construction activity analysis from jobsite video streams [Article]. *Journal of Construction Engineering and Management*, 141(11), 19. doi:10.1061/(asce)co.1943-7862.0001010

Lu, W., Huang, G. Q., and Li, H. (2011). Scenarios for applying RFID technology in construction project management. *Automation in Construction*, 20(2), 101-106. doi: 10.1016/j.autcon.2010.09.007

Neve, H. H., Lerche, J., and Wandahl, S. (2021). Lean construction methods' impact on construction labour efficiency. *Proceedings of IGLC29*, Lima, Peru.

Neve, H. H. and Wandahl, S. (2018). Towards identifying making-do as lead waste in refurbishment projects. *Proceedings of IGLC26*, Chennai, India.

Neve, H. H., Wandahl, S., Lindhard, S., Teizer, J., and Lerche, J. (2020a). Determining the relationship between direct work and construction labor productivity in North America: Four decades of insights. *Journal of Construction Engineering and Management*, 146(9), 04020110. doi: 10.1061/(ASCE)CO.1943-7862.0001887

Neve, H. H., Wandahl, S., Lindhard, S., Teizer, J., and Lerche, J. (2020b). Learning to see value-adding and non-value-adding work time in renovation production systems. *Production Planning & Control*. doi: 10.1080/09537287.2020.1843730

Ohno, T. (1988). *Toyota Production System*. Productivity Press.

Olivieri, H., Seppänen, O., and Peltokorpi, A. (2017). Real-time tracking of production control: requirements and solutions. *Proceedings of IGLC25*, Heraklion, Greece.

Pérez, C. T., Salling, S., and Wandahl, S. (2022). Five guidelines for adopting smartwatches in construction: a novel approach for understanding workers' efficiency based on travelled distances and locations. *Sustainability*, 14(14), 8875. doi: 10.3390/su14148875

Pérez, C. T., Sommer, L., Costa, D. B., and Formoso, C. T. (2015). A case study on causes and consequences of transportation waste. *Proceedings of IGLC23*, Perth, Australia.

Picard, H. E. (2002). Construction process measurement and improvement. *Proceedings of IGLC10*, Gramado, Brazil

Ramaswamy, K. P. K. and Satyanarayana N. (2009). Waste in Indian building construction projects. *Proceedings of IGLC17*, Taipei, Taiwan.

Ryu, J., Seo, J., Jebelli, H., and Lee, S. (2018). Automated action recognition using an accelerometer-embedded wristband-type activity tracker. *Journal of Construction Engineering and Management.*, 145(1). doi: 10.1061/(ASCE)CO.1943-7862.0001579

Shahtaheri, M., Nasir, H., and Haas, C. T. (2015). Setting baseline rates for on-site work categories in the construction industry. *Journal of Construction Engineering and Management*, 141(5). doi: 10.1061/(ASCE)CO.1943-7862.0000959

Siriwardana, C. S. A., Titov, R., and Ruwanpura, J. (2017). A study to investigate the relationships between work sampling categories and productivity for concrete slab formwork. *Proceedings of 8th International Conference on Structural Engineering and Construction Management*, Sydney, Australia.

Teizer, J., Neve, H., Li, H., Wandahl, S., König, J., Ochner, B., König, M., and Lerche, J. (2020). Construction resource efficiency improvement by Long Range Wide Area Network tracking and monitoring. *Automation in Construction*, 116, 103245. doi: 10.1016/j.autcon.2020.103245

Thomas, H. R. (1991). Labor productivity and work sampling: the bottom line. *Journal of Construction Engineering and Management*, 117(3), 423-444. doi: 10.1061/(ASCE)0733-9364(1991)117:3(423)

Thompson, S. K. (1987). Sample size for estimating multinomial proportions. *The American Statistician*, 41(1), 42–46.

Tzortzopoulos, P., Kagioglou, M., and Koskela, L. (2020). *Lean Construction – Core Concepts and New Frontiers*. Routledge.

Wandahl, S., Neve, H. H., and Lerche, J. (2021). What a waste of time. *Proceedings of IGLC29*, Lima, Peru.

Wandahl, S., Pérez, C. T., Salling, S., and Lerche, J. (2022). Robustness of work sampling for measuring time waste. *Proceedings of IGLC30*, Edmonton, Canada.

Wernicke, B., Lidelöw, H., and Stehn, L. (2017). Flow and Resource Efficiency Measurement Method in Off-Site Production. *Procedings of IGLC25*, Heraklion, Greece.

Yin, R. K. (2003). *Case Study Research - Design and Methods* (3ed.). Sage Publications, Inc.

Zhao, J., Seppänen, O., Peltokorpi, A., Badihi, B., and Olivieiri, H. (2019). Real-time resource tracking for analyzing value-adding time in construction. *Automation in Construction*, 104, 52-65. doi: 10.1016/j.autcon.2019.04.003

Søren Wandahl, Professor, PhD, M.Sc.ENG, is a full Professor and deputy head of department at Dept. of Civil and Architectural Engineering at Aarhus University. His research focuses on production planning and control and construction labour productivity.

Cristina Toca Pérez, PhD, M.Sc. is a postdoctural researcher at Dept. of Civil and Architectural Engineering at Aarhus University. Her doctoral research, in 2018, was reward the best research project applied in the Brazilian Construction Industry. Her research focuses on Construction 4.0, and Lean construction.

Stephanie Salling, M.Sc. is a research assistant at Dept. of Civil and Architectural Engineering at Aarhus University. She has experience from the construction industry as site manager. Her research focuses on Lean construction and sustainability.

Hasse Højgaard Neve, PhD, M.Sc. is a Management Consultant working at PwC and has a deep interest in, and goal of, optimizing productivity in the construction industry. He works towards this by helping construction industry clients apply the newest knowlegde from both industry and academia.

**Appendix 1**

A table showing the aggregated data points for all N=207 can be found on this URL:

https://www.dropbox.com/s/2ndyouo97yak6sk/Appendix1.pdf?dl=0

**Appendix 2**

A table showing each case's ability to predict other cases, N=828, can be found on this URL:

https://www.dropbox.com/s/raoixr5j4cyjv7i/Appendix2.pdf?dl=0