# Efficiency improvement by navigated safety inspection involving visual clutter based on the random search model

Xinlu Sun[a]; Heap-Yih Chong[b]; Pin-Chao Liao[a*];

[a]*Department of Construction Management, Tsinghua University, Beijing, China*

[b]*School of Built Environment, Curtin University, Perth, Australia*

Xinlu Sun, Research Assistant, Department of Construction Management, Tsinghua University, No. 30, Shuangqing Rd., HaiDian District, Beijing 100084, P.R. China. E-mail: sunxl-2005@163.com

Heap-Yih Chong, Senior Lecturer, School of Built Environment, Curtin University, GPO Box U1987, Perth WA6845, Australia. E-mail: heap-yih.chong@curtin.edu.au

[*]Corresponding author:

Pin-Chao Liao, Associate Professor, Department of Construction Management, Tsinghua University, No. 30, Shuangqing Rd., HaiDian District, Beijing 100084, P.R. China. E-mail: pinchao@tsinghua.edu.cn

Acknowledgement

We thank the National Natural Science Foundation of China (No. 51578317) and the United Technologies Corporation (No.20153000259) for their support of this study.

# Efficiency improvement by navigated safety inspection involving visual clutter based on the random search model

Navigated inspection seeks to improve hazard identification (HI) accuracy. With tight inspection schedule, HI also requires efficiency. However, lacking quantification of HI efficiency, navigated inspection strategies cannot be comprehensively assessed. This work aims to determine inspection efficiency in navigated safety inspection, controlling for the HI accuracy. Based on a cognitive method of the random search model (RSM), an experiment was conducted to observe the HI efficiency in navigation, for a variety of visual clutter (VC) scenarios, while using eye-tracking devices to record the search process and analyze the search performance. The results show that the RSM is an appropriate instrument, and VC serves as a hazard classifier for navigation inspection in improving inspection efficiency. This suggests a new and effective solution for addressing the low accuracy and efficiency of manual inspection through navigated inspection involving VC and the RSM. It also provides insights into the inspectors' safety inspection ability.

Key words: Random search model; navigated inspection; visual clutter; inspection efficiency; construction safety; safety management.

## 1. Introduction

Safety inspection is one of the most important aspects in construction safety management. However, the low efficiency of human visual inspection remains a common issue in construction sites. Manual inspection tasks are time-consuming, resource-hungry, and inefficient. This is especially true for industrial sites and large contemporary systems such as airports, with safety issues such as cracks on concrete surfaces and improper use of personal protective equipment (PPE) [1-6]. To improve the inspection efficiency, automated approaches are desired For example, Dong et al. [5] used pressure sensors and positioning methods to assess the proper wearing of PPE. Li et al. [3] combined image processing and pattern matching algorithms for micro-accessory quality detection. Cheng et al. [4] focused on particular features, such as gray value features and contour features of cracks on the bridges' surfaces, and subsequently developed software for detection of cracks using image-processing methods. Despite the growing efforts in the development of robotic inspection systems, the current applications of state-of-the-art technologies do not guarantee aversion of safety risks at construction sites [5]. In addition, present automated inspection methods are mostly suitable for detection of single and specific targets. Paradoxically, with enormous amount of uncertainty, construction safety inspection cannot be adequately accomplished by robotic systems to completely replace human inspection. Hence, manual inspection is still necessary for some knowledge-intensive and domain-specific tasks [6]. Regretfully, most research focuses on manual inspection studies for enhancement of detection accuracy. However, a contradiction may exist as efficiency and accuracy usually cannot be guaranteed simultaneously [1], introducing a bottleneck for the improvement of both inspection accuracy and efficiency.

Hence, this research aims to explore and determine the levels of inspection efficiency through navigated inspection for construction safety inspection. This research adopted a novel classifier for risk ranking, namely, visual clutter (VC), to reduce the memory workload for inspectors. In addition, a cognitive method of the random search model (RSM) was used to measure the inspectors' inspection efficiency. Ultimately, this study is likely to contribute and extend the knowledge of navigated inspection for construction safety management.

## 2. Literature review

### *2.1 The development of navigated inspection vs manual inspection*

To improve the safety at construction sites, researchers have struggled to improve inspection performance, using a variety of methods. Some studies considered invalid hazard detection to result from human cognitive failure. These studies strongly accounted for human errors and sought more effective safety management measurements from the perspective of human behavior [7-11]. Woodcock [7] established a model for the process of safety inspection in several domains, including amusement ride inspection, food inspection, and construction workplace inspection. In this model, he asserted that individual inspectors with different levels of knowledge and experience conduct inspection tasks in different ways. Perlman et al. [8] explored the hazard recognition skill discrepancy using two different training methods: 1) training by photographs and 2) training by virtual reality. Their experiment demonstrated that correct risk detection increased for those superintendents that were trained on a virtual construction environment; this was attributed to improved cognitive learning. Anu et al.

[12] argued that the previous hazard taxonomy according to the symptoms of errors is inefficient and extremely time-consuming. They proposed that human error taxonomy should refer to the cause of problems, for improving software inspection performance. Others proposed some supplemental instruments for safety inspection, such as dynamic risk taxonomy [13], a real-time performance feedback system [14], and mobile computing methods [15]. The risk items were classified according to frequency or severity.

Recently, more advanced technologies have been developed, such as automated inspection approaches that use artificial intelligence and computer vision methods. Moczulski [6] suggested that robotized inspection and diagnostics could replace human experts. He considered several issues concerning robotized inspection and introduced a few application examples, such as inspection of underground galleries, wind turbines, and aircrafts. Rea et al. [2] designed a robotic system for industrial sites' inspection and monitoring. In their system, a method of three-dimensional (3D) mapping was utilized to reconstruct objects, and computer vision methods were used for extracting defects. However, the developed model remains a prototype model and requires further development. Dong et al. [5] combined building information modeling (BIM), real-time location systems (RTLS), and sensors, to alert workers themselves and safety managers about improper use of PPE. BIM and RTLS were used for location tracking and for deciding on the necessity of using PPE, while pressure sensors indicated whether the

PPE was used properly. Furthermore, Li et al. [3] also utilized an image-processing method for micro-accessory quality inspection.

Although it seems that automated inspection systems are preferred in various industrial domains as well as at construction sites, manual safety inspection is still irreplaceable at construction sites, due to the following reasons. First, in spite of the excellent data mining capability of state-of-art artificial intelligence methods, human inspection on knowledge-intensive tasks remains necessary [6]. In addition, robotic systems can do dangerous tasks, but decision-making still needs to be made by humans [2]. Overall, construction sites are full of dynamic and complex activities, usually involve a large number of risk issues. Obviously, it is a huge project for an automated robotic system to learn and identify all risk items or hazards. Nevertheless, manual inspection suffers from much 4criticism because it is resource-hungry, time-consuming, and not efficient [1, 2, 6]. To overcome these drawbacks, this paper posits that navigated inspection equipped with a task-oriented daily checklist can significantly improve manual inspection performance, hazard detection accuracy, and detection efficiency.

#### *2.1.1 The need for objective and quantitative assessment of navigated inspection*

First and foremost, it is necessary to use an objective and quantitative assessment of the inspectors' visual search performance. However, present measurements utilized for inspection performance are not well established. Interviews and feedback collection are commonly used to illustrate the potential ability of the proposed system for improving inspection capacity. Zhang et al. [15] interviewed several experienced experts in construction industry, and the interviewed experts acclaimed the prototype safety management tools used in mobile safety management applications. Anu et al. [12] also asked their study subjects to provide feedback of the designed human error taxonomy and the result proved that the authors' idea received excellent users' reviews. Despite the validity of these approaches, these assessments are still relatively subjective and full of uncertainties. A method for objective and quantitative measurements is strongly needed. There are two common quantitative assessments: 1) the number of correct answers and 2) the ratio of the number of correct answers to the time consumed. Perlman et al. [8] examined risk detection accuracy of inspectors trained on paper documents or on a virtual environment, and they found that training on the virtual environment increased the number of correctly detected risks. Similarly, Anu et al. [12] analyzed the two indexes of effectiveness and efficiency as the quantitative assessment for software inspection. Briefly, effectiveness refers to the number of identified targets, while efficiency refers to the ratio of the number of identified targets to the time consumed. Undoubtedly, an index demonstrating time-variant inspection performance is critical for assessing inspection efficiency. However, the detection rate (DR) obtained by dividing the number of correct answers by time leads to the loss of individual-related information, as this method obtains the average level of the group.

Researchers in the field of psychology prefer to analyze eye movement characteristics for quantifying visual search performance. The most representative indexes are fixation time percentage in the area of interest (AOI), fixation count percentage in AOI, response time, and duration time of each fixation [16-18]. The indexes reflect various processes of visual search. For instance, fixation time percentage in AOI demonstrates the percentage of attention allocated to the target rather than to the background, while duration time of each fixation highlights the speed of information processing and decision making. Therefore, eye movement characterization is also inappropriate for assessing inspection efficiency. An integrated measurement that can make use of individual data and delineate the inspection capacity would be more suitable for safety inspection.

#### *2.1.2 RSM for inspection efficiency*

To quantitatively and objectively assess inspection efficiency, RSM can be considered. Manual visual inspection is believed to be a cognitive behavior called visual search [19]. Conventionally, a visual search task should be performed under two common strategies, namely, systematic search strategy and random search strategy [20-22]. In the random search strategy, the observer's fixations randomly address the entire visual scene, whereas in the systematic search approach, a certain fixation never overlaps with previous fixations. This indicates that in the random search approach, an observer may focus on a certain object repeatedly in the stimulus scene. Actually, the systematic search model is based on the hypothesis of perfect memory, while the RSM is based on the hypothesis of absolutely imperfect memory. Therefore, an actual visual search process is never conducted absolutely and exclusively in either search mode; rather, it combines the two strategies. Visual search capacity can be defined in terms of the percentage of targets detected against the time consumed. Theoretically, the RSM is captured by an exponential curve, while the systematic model is captured by a linear function [19, 21]. In addition, when the background area is larger than the search target, observers tend to follow the random search strategy [20, 21, 23]. Many practical visual search tasks, such as baggage inspection using X-ray technology, industrial quality or safety inspection, fall in this category.

Apart from that, previous studies on the RSM mainly focused on exploring the performance of the RSM in various scenarios. Yu et al. [19] found that the traditional RSM performed very well on static visual search tasks under dynamic conditions. Chan et al. [20] used the RSM for a double-target search task and the results showed that the RSM fitted both the individual and pooled data very well. These explorations explain that most people choose the random search strategy for most practical inspection tasks. However, these are relatively limited theoretical studies on the application of the RSM to pragmatic inspection tasks. Hence, this research aims to bridge the existing gap and attempts to use the RSM to measure inspection capacity on construction safety inspection under the random search strategy.

### ***2.2 Factors affecting navigated inspection***

#### *2.2.1 Experience and working memory are the two critical factors that affect inspection efficiency*

As demonstrated, manual inspection is irreplaceable, even under a fully navigated inspection. To prove that safety hazard inspection can be streamlined via navigation, it is necessary first to identify the pertinent controlled variables. In reality, there are two common factors affecting safety inspection, namely, experience and working memory.

Undeniably, experience is one of the most pivotal factors that reflects an inspector's professional aptitude. The strategy of employing highly experienced inspectors is a common and effective approach for obtaining more reliable hazard-detection results. Woodcock [7] proved that experienced inspectors conducted search tasks differently, compared with inexperienced ones. Experienced inspectors can solve most of the arising issues themselves by analyzing similar scenarios or turning uncertainties into manageable conditions. Psychologists also explicitly illustrated how experience improves the search performance. Nodine et al. [24] also reiterated that experts spend less time on search compared with novices. Schyns et al. [25] asserted that experts perform better because they have substantial knowledge of the situation. Essentially, experience-related knowledge is accumulated and stored in the long-term memory [26, 27]. Correspondingly, the working memory affected by navigation is another critical factor for safety inspection.

During visual search, the human visual system receives much more information than it can process, necessitating the attentional mechanism called selective attention to filter and select only the information that is useful for the process. Navigated inspection offers search templates of hazards in the inspector's working memory by looking through a checklist before inspection, serving as a top-down guidance for selective attention [28, 29]. The guidance from the working memory can quickly focus selective attention on relevant targets rather than background images [30, 31]. On the other hand, if background distractors match the templates, visual search will take longer to accomplish. In addition, search templates facilitate information processing to verify that the suspected item is a hazard. In addition, the width and precision of target templates in the working memory significantly affect the decision time to verify the targets [26].

#### *2.2.2 VC serves as a potential classifier for navigated inspection*

The search templates, obtained from navigation, serve as a top-down guidance. Yet, there is another guidance mechanism, which relates to looking at the outstanding features of targets, which can be identified from differences in luminance, color, motion, orientation, or size between items [32, 33]. It was shown that a contrast in at least one dimension between an object and its background can capture an observer's attention [29]. The search ability is affected by both top-down and bottom-up patterns [34]. Toet [35] demonstrated that the energy contrast and structural dissimilarity between targets and distractors affect the perceptual and conceptual search performance respectively.

Hence, VC can be adopted to measure salient targets [16, 27, 36, 37]. VC is defined as background images that confuse and distract observers. Schmieder et al. [36] illustrated that the searching ability is negatively affected by high VC. Boersema et al. [38] concluded that increasing the number of fixations increased the search time in the presence of high VC. These studies failed to create the best experience of human machine interaction. Ji et al. [16] examined the pedestrian assistant efficacy for two types of night-vision enhancement systems, evoking different levels of VC. Liggins et al. [39] evaluated the effectiveness of color-display night vision goggles against a monochromatic night-vision scene background. Recently, researchers also analyzed synergistic effects of VC with other factors, such as the aging effects on the apperceived VC [40], and an integrated measure of display clutter based on feature content, user knowledge, and search performance [41].

Although the effect of VC on search capacity is relatively unambiguous, the moderation effect of VC on the relationship between working memory and search efficiency has not been sufficiently explored. VC could be a potential classifier for navigated inspection for classifying risk items and for shortening the risk checklist [29]. These previous studies verified that risk detection accuracy in navigation varies with clutter, but did not account for time. In this research, we intend to explore whether navigated inspection that uses VC as a classifier can improve risk detection efficiency.

### *2.3 Development of Hypotheses*

Manual inspection is irreplaceable, owing to the deficiency of automated inspection methods. However, manual inspection suffers from low accuracy and efficiency of safety inspection. This paper argues that navigated inspection that uses VC as a risk classifier can increase inspection accuracy and efficiency. The search efficiency of dependent variables can be measured using the RSM, while the independent variables are the working memory and VC. Moreover, the inspectors' experience is controlled as an important factor that determines the search ability. Consequently, this research proposes two hypotheses, namely

$H_0$: The RSM can be utilized to measure hazard detection efficiency.

$H_1$: Navigated inspection affects hazard detection differently for different levels of VC.

## 3. Methodology

### *3.1 Factors and the RSM model*

#### *3.1.1 Factors*

Factorial approach was used in the present study for experimental design. The independent variables were the existence of search templates, VC, and search time. The dependent variable was the cumulative probability of risk detection. In addition, the inspectors' experience should be consistent.

The existence of search templates represents the discrepancy of the working memory. In this experiment, the search templates of hazards were offered by navigation, which meant prior stimuli of critical hazards. The study participants were divided into the experimental and controlled groups, according to the received prior stimuli. This indicated that the participants in the experimental group would have the search templates of critical hazards, while the participants in the controlled group would not have these search templates.

VC was a moderation variable in this experiment. The VC of the stimuli images featured natural construction hazards, capturing salient targets. Four indexes (color, size, distinction, and orientation), were considered and incorporated into VC. Six basic classifiers were used to calculate the four indexes: object category number, number of brilliant objects, number of salient objects, number of indistinguishable objects, and number of horizontal objects. For a detailed description of the computation process, the interested reader is referred to the methodology of Liao et al. [29]. The photographs of natural construction sites used in this experiment were all acquired from the Otis Elevator Company. Table 1 shows the levels of image VC, for the fifteen considered photographs.

<Table 1 near here.>

Figures 1a-1c are photographs No.2, 10, 15, from groups of low, median and high VC, respectively.

<Figure 1a near here.>

<Figure 1b near here.>

<Figure 1c near here.>

The participants both in the experimental and controlled groups searched hazards in the same fifteen images. The participants in the experimental group received advance knowledge of critical risks, as shown in low-clutter images Nos. 2 and 4, median-clutter images Nos. 6 and 10, and high-clutter images Nos. 11 and 13.

As for the control of experience differences, experience was measured quantitatively, as follows. The ratio of one's working year to the largest one among all the participants (RYII) and familiarity with the fatal prevention audit ( FPA ) checklist (FFPA) were used for measuring the participants' experience. The FFPA used the score of several questions (e.g., how many risk items are in the checklist?) to test the participants' familiarity with the checklist. The final experience score was the mean of RYII and FFPA. A t test on the scores illustrates the consistent experience of the participants in the two groups.

### *3.1.2 The random visual search model for safety inspection*

The general equation of the RSM, given below, suggests the relationship between the cumulative probability of risk detection F(t) and the search time (ST). It can be written as follows:

$$\ln\left(1 - \mathrm{F(t)}\right) = \mathrm{k} - \lambda \mathrm{t} \qquad (1)$$

Here, F(t) refers to the cumulative percentage of risk found within time ST, which also represents the probability of risk detection for an average participant at that time; k and $\lambda$ are the parameters to be estimated; t is the search time of a certain task. The cumulative percentage of risk found within a certain time ST (i.e., F(t)) is given by the proportion of risk items precisely detected for that time.

As for the search time of target detection, researchers previously relied on certain assistant software to guide participants to perform a search task on pictures containing targets and distractions [19, 20]. By registering the left button clicks of a mouse when starting or stopping a task, the software recorded the start and end time of the search process. The difference between the end and start times is the search time. However, errors may occur owing to the interval time that may exist during the action of shifting the subject’s attention to clicking on the mouse button. Noticeably, a more accurate method for measuring the search time, which focuses on the search task, is desperately desired. Researchers in the cognitive psychology domain prefer to use eye-tracking equipment, a device that assists in providing a spatial and temporal record of eye movement characteristics [16, 37, 41]. It has been shown that eye movement characteristics reliably capture different modes of visual processing [24]. Parameters such as fixation duration could be utilized for analysis of visual search processes. In this experiment, an eye-tracking device was used for recording the participants’ search time. Before a picture appeared, the participants were asked to focus on the screen. The time at which the picture appeared on the screen was considered as the beginning of the search process. The participants could stop anytime when they finished searching for a risk item and then divert their attention to the staff and give their decision. The time of the last fixation on the picture denoted the end time of a search task. The eye-tracking device and the assistant software marked the show time of the picture and the last fixation time, and the difference between the two was considered as the search time in the present approach.

## *3.2 Experimental design*

### *3.2.1 Participants and Grouping*

The participants were all male staff from the Otis Company in Beijing or Shanghai, China. Random sampling from a limited set of available subjects was performed and overall 42 participants were selected. All of the study participants were familiar with elevator installation as safety officers, inspectors, debuggers or maintenance workers. The participants were asked to attend a 6-day safety training course taught by the company each year, and they passed the annual safety knowledge examination and safety performance assessment. Hence, the possible impact of job differences can be ignored. In addition, the hazards involved in the experiment were all general hazards that occurred during the installation process. The study participants all had normal or corrected-to-normal visual acuities and no one had dyslexia.

In addition, working memory and experience are two important factors that affected visual search ability in the experiment. Thus, it was necessary to balance the experience and memory parameters between the experimental and control groups.

To eliminate the effects stemming from individual working memory differences, a preliminary experiment was conducted using the Wechsler memory scale (WMS), which is widely used to test multidimensional memory capacity The WMS is typically used for classifying people and for identifying those with memory impairments. The WMS consists of seven subsets and it was revised in 1987 as the Wechsler memory scale-revised (WMS-R), which now accounts for four main memory functions: 1) attention/concentration, 2) verbal memory, 3) visual memory, and 4) delayed memory. Estimating one's entire memory function is impossible. Moreover, it is generally accepted that memory is mediated by complex neuronal networks that are located in different brain regions. Thereupon, only visual memory was tested in the present study. Two indexes were employed in the present study: figural memory and visual reproduction. In the preliminary experiment, the study participants first observed four images of different construction scenes with hazards and tried to retain them. Then, they were asked to recognize the preceding hazards in other eight images and match them with those in the previous images. In this step, correct recognition and matching each rewarded the participants with the score of 0.5. An absolutely correct answer yielded the score of 1. Thus, theoretically, the participants' scores in the preliminary experiment ranged from 0.0 to 8.0. To guarantee both sample size and concordant visual memory ability, only those participants who scored above 2.5 were selected for the formal experiment. Finally, 30 participants, scoring a mean of 4.217 with a standard deviation of 1.023, were randomly divided into the experimental and control groups.

A t test was performed and demonstrated that the variances and means of the participants' experience in the two groups were statistically equal ($p$ = 0.147 and 0.833, respectively). Therefore, the experiment was considered to be well-controlled from the experience viewpoint.

### *3.2.2 Materials*

The main experimental instrument was the eye-tracking device, an SMI iView X™ headset eye tracker (SensoMotoric Instruments, German) assisted by the software Begaze version 3.2, for search time analysis. SPSS version 21 was utilized for compiling and analyzing the experimental data. PowerPoint files were shown on a 19” laptop, helping to present the FPA checklist and images with construction scenes in which risks had to be detected. In total, fifteen pictures were divided into three subsets according to the VC values (high, medium, and low clutter). In the experimental group, each subset consisted of five pictures, two that featured risks for which the participants were well-trained before the search task, while the remaining three pictures featured risks for which the participants were not trained. For the control group, the participants were trained on neither of the risk items in pictures. After the primary searching process in the experiment, paper-formed chromophotographs and FPA checklists were provided to the participants and they were asked to detect and report risk items. All photographs that contained images of construction scenes were supplied by the Otis Company.

### *3.2.3 Procedures*

The experiment encompassed three sections: 1) brief introduction of the experiment and practice, 2) risk searching task, and 3) risk identification answer. The details of the experiment are described below:

(1) Practice and checklist learning (5 + 15 min):

For the initial 5 min, the experimental procedures were explained to all participants, and all participants practiced hazard search without eye-tracking equipment. Then, the participants in the experimental group spent 10 min to review all of the 90 risk items in the FPA checklist. After that, they had 5 min to learn about and try to remember the six critical risk items to obtain search templates in their working memory. Whereas the participants in the control group did not need search templates and spent all of the 15 min reviewing the FPA checklist.

(2) Risk searching (20 min):

After calibrating the eye-tracking device at the beginning of the experiment, the participants in both groups were asked to detect up to one risk item in each image showing a natural construction scene. Fifteen pictures were exhibited one by one. The participants performed the hazard detection task and provided yes/no answers, corresponding to the existence or non-existence of a risk issue. Most of the participants finished all of the fifteen tasks in 20 min. No time limitation was set to eliminate possible psychological pressure on the participants, allowing them to conduct the searching task unaffectedly.

(3) Risk identification (20 min):

In this step, the participants were asked to match the detected risk items with the expressions in the FPA checklist. The checklist and chromophotographs in which the participants declared the existence of a safety hazard were provided. They were expected to write the picture number right beside the matching risk item in the checklist.

## 4. Results

### *4.1 Descriptive statistics*

The age of the 30 participants ranged from 22 to 58 years. In addition, the average time the participants worked in the construction industry was 10 years, with the average of 5 in safety-related positions. Although the work experience time (work age) seemed to vary across the participants, their experience was concordant within the experimental and control groups.

The participants came from the Beijing and Shanghai branches of the Otis Company, with 56.7% from Beijing and 43.3% from Shanghai. As for the education, thirteen of the study participants attended college or university, while the rest held a junior, senior middle school, or a special secondary school diploma. However, only 30% of the participants took an elevator installation relevant major, such as mechanics and engineering supervision. Most of them were debuggers and maintenance workers, and others were safety officers or supervisors.

### *4.2 Risk detection performance*

By referring to the RSM theory, only correct detection responses were analyzed in this experiment. A total of 179 (39.78%) correct responses were obtained in this study. Table 2 shows the correct detection data, for different levels of VC. Generally, the correct

DRs ranged from 24.00% to 50.67%. The mean of the DR for the experimental group was higher than that for the control group (43.11% and 36.44% respectively). Considering the effect of VC, DR was higher for medium and high clutter scenes for the experimental group compared with the control group. However, with decreasing VC, the DR decreased rather than increased in the experimental group, contrary to a general expectation. Figure 2 shows the median search time. Generally, search time increased as VC increased, both for the experimental and control groups. Moreover, the inspectors searched faster with navigation in the experimental group and the effect of search templates on the search time was more significant for higher VC.

<Table 2 near here>

<Figure 2 near here>

Overall, the detection accuracy significantly increased both in high and middle clutter scenes with search templates in working memory.

### *4.3 Validation of the RSM for safety inspection*

Figure 3 shows the cumulative percentage of detection F(t) against search time for different VC levels, for the experimental and control groups. All of the six curves increased exponentially with time, indicating that the visual search mode of the risk detection task could be well described using the RSM. Regression analysis also suggested a good fit. Table 2 lists the estimated regression parameters. The fitting coefficients ($R^2$) were in the 0.821–0.985 range ($p$ < 0.001), demonstrating a high degree of correspondence between the experimental data and the fitted equation. All estimated parameters were statistically significant at the 1% or 10% level. These results helped to confirm that a random search strategy was employed during risk detection under various VC levels. Moreover, from the statistics viewpoint, $\lambda$ in the equation is the reciprocal of the average search time while $\frac{k}{\lambda}$ is the response time ($t_r$). Therefore, combining the two, we obtain the theoretical mean ST ($\frac{1}{\lambda}+t_r$) and median ST ($\frac{\ln 2}{\lambda}+t_r$).

<Figure 3 near here>

<Table 3 near here>

Table 3 shows a comparison between the theoretical and estimated means and median ST. The theoretical mean and the median ST are consistent with the experimental values to some degree. By referring to the mean ST, the theoretical mean was relatively close to the experimental mean in the experimental group (Deviation < 7%), instead of the control group (Deviation > 15%). Regarding the median ST, the discreteness of the deviation was more conspicuous. Deviation in the control group for low and high clutter was relatively large (> 15%), but was quite small for other conditions (< 8%). By and large, the deviation between the theoretical and experimental values of ST was acceptable. This result serves as another proof that the RSM is adequate and valid for describing the visual search mode participants employed in the risk detection tasks. In addition, this conclusion further confirms the validity of the practical application of the RSM in construction safety management. It can help to estimate the effective search time and can serve as a reference for safety inspection planning.

<Table 4 near here>

### *4.4 Significant improvement of inspection efficiency under high and median VC*

After validating the RSM for delineating the detection efficiency, the curves of the fitted model were generated for further analysis, and the generated curves are shown in Figure 4. The six curves could be divided into two groups, captured by the solid and dashed lines, representing the inspection efficiencies of the experimental and control groups, respectively. Different colors were used for different VC scenarios, with red, blue, and green colors corresponding to the high, intermediate, and low VC scenarios, respectively.

<Figure 4 near here>

Obviously, the results are promising. First of all, considering the effect of VC and not considering the working memory, the inspection performance of the control group was negatively correlated with VC, with inspection efficiency tending to decrease with increasing VC. This result is in line with the suggestion that VC negatively and significantly affects risk detection performance.

Next, noticeable changes appeared with search templates in the participants' working memory on comparison of the two groups. The curvature increased significantly, indicating that the cumulative risk detection probability was higher for the experimental group compared with the control group during a certain time period. Moreover, the worst inspection performance was better than those registered when no search templates were used, as can be ascertained by analyzing and comparing the red solid curve with the green dashed curve. Based on this, we further conclude that the effect of search templates on inspection efficiency can overshadow the negative effect of VC. The random visual search mode tended to be more close to the systematic mode with the enhanced working memory. In addition, the discrepancy between the differently colored curves was smaller for the experiment group. This result is in line with the modulation of the VC effect by search templates.

Furthermore, the balance between accuracy and efficiency is key for improving the inspection efficiency. Ideal (100%) accuracy can never be attained. For example, even highly experienced inspectors may detect 80% of risk issues correctly [9]. Therefore, if 80% is the benchmark, ST saved by search templates can be determined. Figure 5 shows the time gain on the tasks.

<Figure 5 near here.>

The crossover between the horizontal line and the different curves represents the search time cost for the DR of 80%. The length of the line segments between two same-colored points corresponds to the ST gain when using the search template, for the corresponding VC. Apparently, the ST gain for low clutter scenes was much smaller than that for high and intermediate clutter scenes. Approximately, search templates reduced the ST by less than 10 s for low VC, while the reduction was 20 s for high and intermediate VC scenarios. This suggests that search templates cannot be efficiently utilized for analysis of low clutter scenes. Simultaneously, the ST at the 80% detection probability with normal working memory and low VC conditions, referring to the green dashed line, is quite close to that for the enhanced working memory and high and intermediate VC conditions. This also reveals that a few distractors in low clutter scenes are not likely to confuse inspectors. Therefore, the detection performance was nearly as good as that for high clutter scenes with search templates in the working memory, alleviating the necessity to use search templates in low clutter scenes. Moreover, it is advisable to exclude risk items in low clutter scenes from safety training, which will likely improve inspection efficiency.

# 5. Discussion

## *5.1 Implications and interpretation of the RSM*

This study has validated the effect of VC on risk detection and RSM in measuring safety inspection efficiency. Let us note first that the cognitive process associated with visual search was affected by VC on safety risk inspection tasks. Abundance of distractors in high clutter scenes impeded the search process and reduced the search efficiency, increasing the time required to complete the task and reducing the accuracy of detection. In addition, this study confirmed the adequacy of using the RSM; consequently, the RSM can be used as a quantifiable framework for measuring safety inspection. The scatter diagram showed that the cumulative probability of risk detection increases exponentially with time. Furthermore, the regression results indicated that the experimental data is very well fitted by the RSM. This confirms that the RSM is adequate for safety inspection efficiency measurements.

In addition, the effect of VC was evaluated not only for normal inspection, but also for navigated inspection, the latter corresponding to the situation in which the study participants were supplied with search templates in their working memory. The results are inspiring. Navigated inspection improved the risk detection accuracy and reduced the time on task; while VC had the opposite effect. The collision of the two factors might bring surprising effects. In this work, these effects were explored, and it was found that the effectiveness of the present navigated inspection depended on VC. For the risk DR of 80%, the highest rate for experienced inspectors, navigated inspection yielded the time gain of 20 s approximately for high and intermediate clutter scenes. Nevertheless, for low clutter scenes, the time gain was less than 10 s for each hazard detection. Moreover, normal inspection without navigation was relatively quick for low clutter scenes. Consequently, navigated inspection should further focus on hazards for high and intermediate clutter scenes and give them priority for achieving efficient inspection.

In addition, the regression results were not predictable. Confusion may exist regarding the parameter estimation and the RSM deviation analysis. For example, the values of $k$ in the control group were all negative. Moreover, the deviations between the experimental mean ST and theoretical mean ST for the control group, as well as the median ST, were relatively large (larger than 15% and up to 23% in most models). Obviously, the maximal STs for the control group were much larger than those for the experimental group. But, the STs were not so large for lower cumulative probabilities.

This indicates that with increasing time on task, the cumulative risk detection probability of normal inspection increases initially as in navigated inspection. However, for detection probabilities above 50%, longer times on task are required. In addition, regarding the initial detection probability in normal inspection, the probability is still strictly positive even when the participants did not know the correct answer and had to guess. Without any search templates in the working memory, inspectors had to recall from their long-term memory and they had to consider all of the 91 risk items according to the checklist. This process might be confusing and the participants instead might have opted to guess. That could explain the unexpected negative values of $k$ and the observed deviation between the experimental and theoretical statistical values.

### *5.2 VC affects search performance differently*

A general agreement exists that background distractors are likely to confuse observers and impair search performance. However, there is a conflict between distracting and informative non-target objects. On the one hand, distractors can indeed bother observers. Ho et al. [34] found that search time was longer for high clutter scenes. Ji et al. (2010) conducted an experiment to examine the drivers' performance on the pedestrian detection task, for differently cluttered scenes, using night vision assistant devices. The results suggested that high VC impairs pedestrian detection, increasing the time on task and reducing the detection success probability. These results confirm that VC impairs detection. On the other hand, non-target objects are only distractors. They may convey critical information even though they are not targets themselves [37, 39, 42]. The performance on the risk detection task was negatively affected by VC. In particular, the inspectors detect hazards faster and more accurately when VC was higher. Compared with the results of previous research, this may suggest that background objects at construction sites offer little information for risk identification and act as distractors on such search tasks. They may help for scene perception; however, limited relationships between distractors and hazards were uncovered.

### *5.3 Interpretation on the mechanism of VC involved navigation*

Navigated inspection can significantly improve risk detection efficiency for high and intermediate clutter scenes, compared with normal inspection. On the other hand, for low clutter scenes, the performance of normal inspection was relatively good, and its improvement brought about by navigation was quite limited. This can be explained using the framework of the feature integration theory.

As explained in the literature review, search templates obtained from navigated inspection are stored in the working memory, which offers top-down guidance on selected attention for quickly focusing on targets. Moreover, selected attention is also affected by bottom-up guidance via salient targets. The feature integration theory explains that these two types of guidance can be separated into two stages [43]. Treisman et al. [43] found that features of objects, such as colors, orientations and shapes, are perceived serially and registered automatically early in the visual search process. Then, at a later stage, these features are located and integrated for identification of objects. Until this stage, features are analyzed and examined for matching the search templates. Without navigation and search templates in the working memory, suspected hazards have to be retrieved from the long-term memory and matched to the identified objects. This process is time-consuming. Therefore, search efficiency will be improved by navigation. However, this interpretation does not consider VC. This is likely adequate for intermediate and high clutter scenes, based on the research findings. The intermediate and high VC represent various distractors in the visual field and hence allow to distinguish targets from distractors. Nonetheless, the situation is different for low VC. Now, scarce distractors do not cause significant confusion. The hazard itself can attract attention through salient targets, yielding quick identification. On the one hand, this process is relatively simple and effective, which explains why without navigation the probability curve for the low clutter scene is close to those obtained for navigated inspection. On the other hand, search templates have less impact on the risk search process, because selective attention is guided primarily by salient targets. Hence, navigated inspection exhibits a weaker effect on inspection efficiency for low VC.

# 6. Conclusions

## *6.1 Contributions*

This research has explored and determined the idea that navigated inspection can significantly improve inspection efficiency. The theoretical and quantitative search efficiency measurements have been validated through the application of the RSM for construction safety inspection. The good correspondence of the experimental data to theoretical results suggests that the present instrument is compelling and powerful for non-subjectively measuring inspection efficiency. It provides new and useful insights into the inspectors' ability to perform safety inspection tasks. Beyond the theoretical and experimental measurements, the RSM can be used for practical safety management. Using software to display hazard scenes and to record the search time, employers can measure inspectors' risk detection ability and design ad-hoc trainings. Moreover, improvement of the inspection ability by training can also be measured utilizing RSM.

Apart from that, the present work has explored the effect of VC on the efficiency of navigated inspection by measuring inspection ability using the RSM. The search efficiency of normal inspection is relatively high and rarely improved by navigation when VC is low. However, for intermediate or high clutter scenes, the risk detection task is eminently time-consuming and navigation using search templates significantly improves the detection efficiency. This result validated the VC as a risk classifier, and can be further used to optimize the effect of navigated inspection.

### *6.2 Limitations and future research*

This study focused on the feasibility of navigated inspection involving VC. Certain limitations need to be considered. The experimental design was confined to indoor space and laboratory conditions. Future studies will assess the experimental method in an outdoor environment. This may yield more precise and convincing results. For practical applications of VC, the integrated measure considering feature content, observers' knowledge, and performance could be explored as it consolidates the primary factors influencing hazard inspection [41].

## References


1. Zhang Y, Zhou Z, Tang K. Sweep scan path planning for five-axis inspection of free-form surfaces. Robot Comput Integr Manuf. 2018;49:335-348.
2. Rea P, Ottaviano E. Design and development of an Inspection Robotic System for indoor applications. Robot Comput Integr Manuf. 2018;49:143-151.
3. Li D, Wang S, Fu Y. Quality detection system and method of micro-accessory based on microscopic vision. Mod Phys Lett B. 2017:1750270.
4. Cheng E, Ming, Zhang GQ, Wang JW. A New Method for Inspecting Crack of Concrete Bridges Using Image Processing Technique. Adv Mat Res. 2010;139-141:2704-2708.
5. Dong S, Li H, Yin Q. Building information modeling in combination with real time location systems and sensors for safety performance enhancement. Saf Sci. 2018;102:226-237.
6. Moczulski W. Robotized Inspection and Diagnostics – Basic Issues. International Conference on Diagnostics of Processes and Systems; 2017 Sep 11-13; Sandomierz, Poland. Cham (Switzerland): Springer; 2018. p. 131-142.
7. Woodcock K. Model of safety inspection. Safety Science. 2014;62(2):145-156.
8. Perlman A, Sacks R, Barak R. Hazard recognition and risk perception in construction. Saf Sci. 2014;64(4):22-31.
9. Saha S, Jeary A. The effects of experience & interruption in predicting the error & performance rate for a construction inspection task. Neural Comput Appl. 2017;21(6):1191-1204.
10. Jagars-Cohen CA, Menches CL, Jangid YK, et al. Priority-Ranking Workload Reduction Strategies to Address Challenges of Transportation Construction Inspection. J Transp Res Rec. 2009;2098(2098):13-17.
11. Liao PC, Ding J, Wang X. Enhancing Cognitive Control for Improvement of Inspection Performance: A Study of Construction Safety. International Conference on Engineering Psychology and Cognitive Ergonomics; 2016 Jul 17-22; Toronto, Canada. Cham (Switzerland): Springer; 2016. p. 311-321.
12. Anu V, Walia G, Hu W, et al., editors. Effectiveness of Human Error Taxonomy during Requirements Inspection: An Empirical Investigation. SEKE; 2016 Jul 1-3; San Francisco, CA. Pittsburgh (PA): Knowledge Systems Institute; 2016. P. 244-254.
13. Zhang H, Chi S. Real-time information support for strategic safety inspection on construction sites. 30th International Symposium on Automation and Robotics in Construction and Mining; 2013 Aug 11-15; Montréal, Canada. London: International Association for Automation and Robotics in Construction; 2013. p. 506-513.

14. Karasinski J, Robinson S, Handley P, et al., editors. Real-Time Performance Feedback in a Manually-Controlled Spacecraft Inspection Task. AIAA Modeling and Simulation Technologies Conference; 2017 Jan 9-13; Grapevine, TX. Reston (VA): American Institute of Aeronautics and Astronautics, 2017. p. 137.
15. Zhang H, Chi S, Yang J, et al. Development of a Safety Inspection Framework on Construction Sites Using Mobile Computing. J Manag Eng. 2016;33(3):04016048.
16. Ji HL, Tsimhoni O, Liu Y. Investigation of Driver Performance With Night Vision and Pedestrian Detection Systems—Part I: Empirical Study on Visual Clutter and Glance Behavior. IEEE trans Intell Transp Syst. 2010;11(3):670-677.
17. Velichkovsky BM. From levels of processing to stratification of cognition. Vopr Psikhol. 1999 (4):58-74.
18. Velichkovsky BM, Rothert A, Kopf M, et al. Towards an express-diagnostics for level of processing and hazard perception. Transp Res Part F Traffic Psychol Behav. 2002;5(2):145-156.
19. Yu RF, Chan AHS. Display Movement Velocity and Dynamic Visual Search Performance. Hum Factor Ergon Manuf Serv Ind. 2014;25(3):269-278.
20. Chan AS, Chan CY. Validating the random search model for a double-target search task. Theor Issues Ergon Sci. 2000;1(2):157-167.
21. Chan AH, Yu R. Validating the random search model for two targets of different difficulty. Percept Mot Skill. 2010;110(1):167.
22. Karwan MH, Morawski TB, Drury CG. Optimum speed of visual inspection using a systematic search strategy. A I I E Trans. 1995;27(3):291-299.
23. Seung-KweonHong, Drury C. Sensitivity and validity of visual search models for multiple targets. Theor Issues Ergon Sci. 2002;3(1):85-110.
24. Nodine CF, Mellothoms C, Kundel HL, et al. Time course of perception and decision making during mammographic interpretation. AJR Am J Roentgenol. 2002;179(4):917-923.
25. Schyns PG, Rodet L. Categorization creates functional features. J Exp Psychol Learn Mem Cogn. 1997;23(3):681-696.
26. Hout MC, Goldinger SD. Target templates: the precision of mental representations affects attentional guidance and decision-making in visual search. Atten Percept Psychophys. 2015;77(1):128-49.
27. Schuster D, Rivera J, Sellers BC, et al. Perceptual training for visual search. Ergonomics. 2013;56(7):1101-15.
28. Desimone R, Duncan J. Neural mechanisms of selective visual attention. Annu Rev Neurosci. 1995;18(1):193-222.
29. Liao P-C, Sun X, Liu M, et al. Influence of visual clutter on the effect of navigated safety inspection: a case study on elevator installation. Int J Occup Saf Ergon. 2017:1-15.

30. Kunar MA, Flusberg S, Wolfe JM. The role of memory and restricted context in repeated visual search. Percept Psychophys. 2008;70(2):314-28.
31. Frings C, Wentura D, Wühr P. On the fate of distractor representations. J Exp Psychol Hum Percept Perform. 2012;38(3):570.
32. Nothdurft HC. Attention shifts to salient targets. Vis Res. 2002;42(10):1287.
33. Theeuwes J. Top-down and bottom-up control of visual selection. Acta Psychol. 2010;135(2):77.
34. Ho G, Scialfa CT, Caird JK, et al. Visual search for traffic signs: the effects of clutter, luminance, and aging. Hum Factor. 2001;43(2):194.
35. Toet A. Matched filtering determines human visual search in natural images. Proc SPIE Int Soc Opt Eng. 2011;8014(3):1-12.
36. Schmieder DE, Weathersby MR. Detection Performance in Clutter with Variable Resolution. IEEE Trans Aerosp Electron Syst. 1983;aes-19(4):622-630.
37. Moacdieh NM, Sarter N. Using Eye Tracking to Detect the Effects of Clutter on Visual Search in Real Time. IEEE Trans Hum Mach Syst. 2017;PP(99):1-7.
38. Boersema T, Zwaga HJ, Adams AS. Conspicuity in realistic scenes: an eye-movement measure. Appl Ergon. 1989;20(4):267-73.
39. Liggins EP, Serle WP. Color Vision in Color Display Night Vision Goggles. Aerosp Med Hum Perform. 2017;88(5):-.
40. Twedell EL, Koutstaal W, Jiang YV. Aging affects the balance between goal-guided and habitual spatial attention. Psychon Bull Rev. 2016:1-7.
41. Jr CP, Kaber DB. An integrated measure of display clutter based on feature content, user knowledge and attention allocation factors. Ergonomics. 2017 (7):1-15.
42. Doyon-Poulin P, Ouellette B, Robert JM, editors. Effects of visual clutter on pilot workload, flight performance and gaze pattern. International Conference on Human-Computer Interaction in Aerospace; 2014 Jul 30-Aug 1; Santa Clara, CA. New York (NY): ACM; 2014:13.
43. Treisman AM, Gelade G. A feature-integration theory of attention. Cogn Psychol. 1980;12(1):97.

List of tables

Table 1. Values of VC

| No. | Group | VC |
|---|---|---|
| 1 | L | 0.5 |
| 2 | | 1.25 |

| | | |
|---|---|---|
| 3 | | 1.3125 |
| 4 | | 1.5 |
| 5 | | 1.5625 |
| 6 | M | 2.39125 |
| 7 | | 3.2 |
| 8 | | 3.375 |
| 9 | | 3.625 |
| 10 | | 3.75 |
| 11 | H | 5.15 |
| 12 | | 5.38 |
| 13 | | 5.5675 |
| 14 | | 6.545 |
| 15 | | 7.5325 |

Note: H = high cluttered; M = median cluttered; L = low cluttered.

Table 2. Correct detection data under different levels of VC

| ***VC*** | **Group** | **Correct detection number** | **Correct detection rate (%)** |
|---|---|---|---|
| **H** | Exp. | 26 | 34.67 |
| | Cont. | 18 | 24.00 |
| **M** | Exp. | 38 | 50.67 |
| | Cont. | 27 | 36.00 |
| **L** | Exp. | 33 | 44.00 |
| | Cont. | 37 | 49.33 |

Table 3. Parameter estimation results of the 6 models.

| ***VC*** | **Group** | $\lambda$ | $k$ | $R^2$ |
|---|---|---|---|---|
| **H** | Exp. | 0.141*** | 0.319*** | 0.985 |
| | Cont. | 0.035*** | -0.279*** | 0.918 |
| **M** | Exp. | 0.153*** | 0.083* | 0.957 |
| | Cont. | 0.043*** | -0.422*** | 0.821 |
| **L** | Exp. | 0.183*** | 0.164*** | 0.967 |
| | Cont. | 0.088*** | -0.216*** | 0.915 |

Note: Cont. = control group; Exp. = experiment group; H = high cluttered; L = low cluttered; M = Medium cluttered; *Significant at the 10% level; ***Significant at the 1% level.

Table 4. A comparison between the theoretical and practical mean and median of ST

| ***VC*** | **Group** | **Experimental mean ST** | **Theoretical mean ST** | **Deviation (%)** | **Experimental median ST** | **Theoretical median ST** | **Deviation (%)** |
|---|---|---|---|---|---|---|---|
| **H** | Exp. | 10.016 | 9.355 | 6.61 | 7.875 | 7.178 | 8.85 |
| | Cont. | 24.360 | 20.600 | 15.44 | 9.658 | 11.833 | 22.52 |
| **M** | Exp. | 7.412 | 7.078 | 4.50 | 5.029 | 5.073 | 0.87 |
| | Cont. | 16.506 | 13.442 | 18.56 | 6.250 | 6.306 | 0.89 |
| **L** | Exp. | 6.533 | 6.361 | 2.64 | 4.550 | 4.684 | 2.94 |
| | Cont. | 11.008 | 8.909 | 19.07 | 4.700 | 5.422 | 15.36 |

List of figures

Figure 1. Means of Search Time

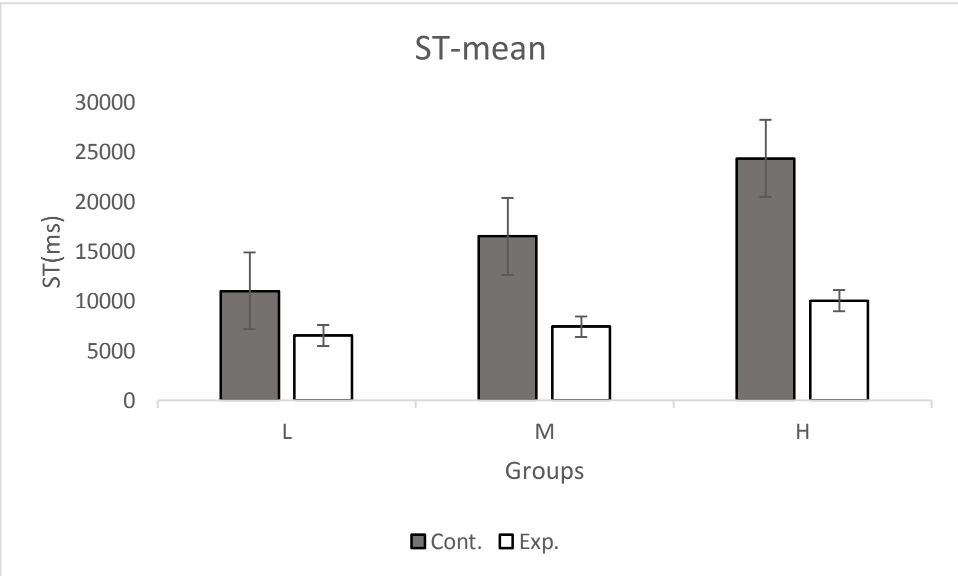



Figure 2.The relationship between ST and F(t)

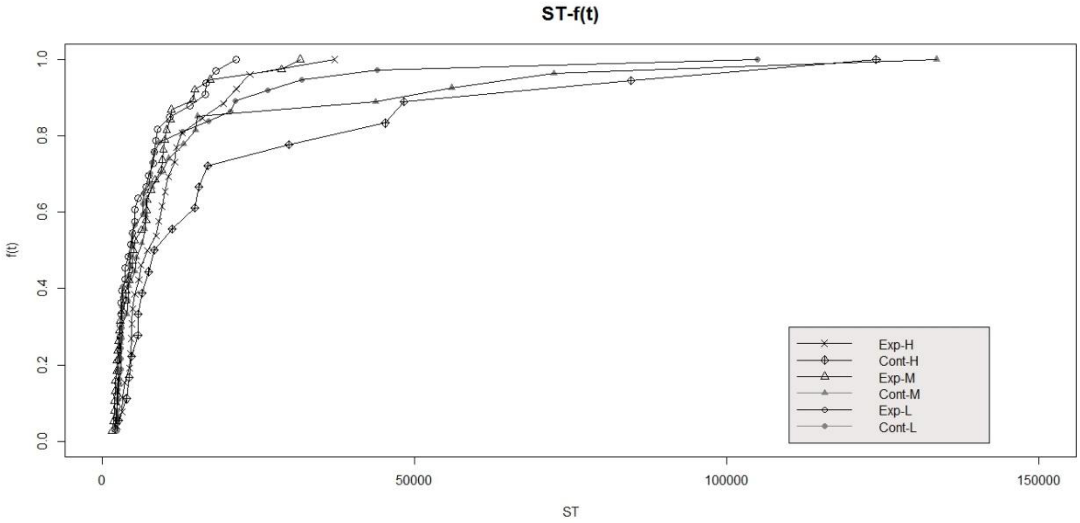



Figure 3. Inspection of RSM under Several Levels of VC

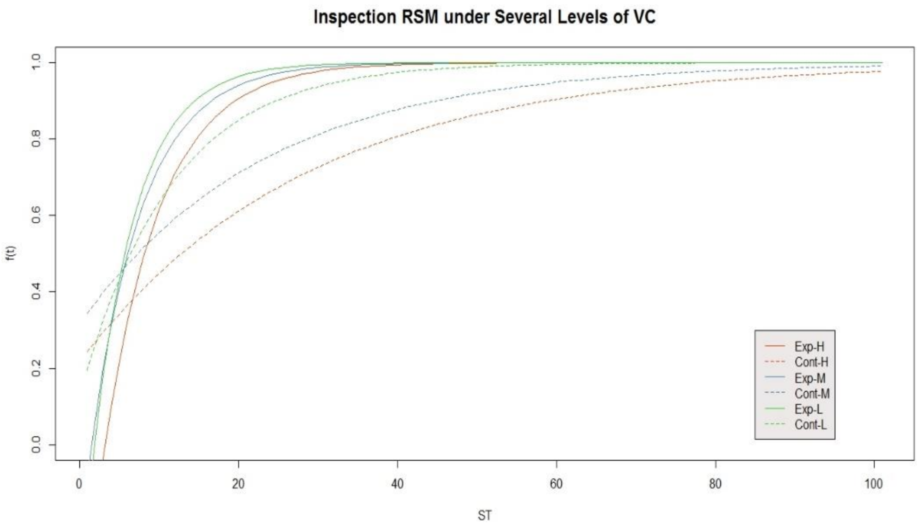
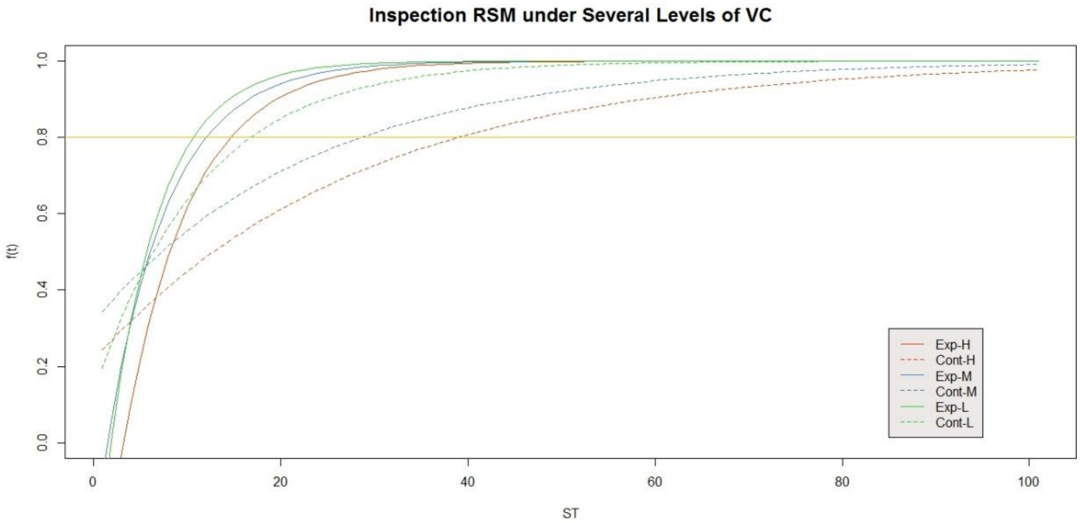



Figure 4: The curves of the fitted model


Inspection RSM under Several Levels of VC
f(t)
ST
Exp-H
Cont-H
Exp-M
Cont-M
Exp-L
Cont-L