Problem Chosen	2025 MCM/ICM Summary Sheet	Team Control Number
A		**2501420**

Treading Through Time: A Wear Tale of Stairs

Summary

In this paper, through a thorough examination of the wear and tear on building staircases, we propose a wear - usage frequency mathematical model for inferring various information about staircases and the human activities involving their use. Additionally, we present an evaluation strategy for some inferences made by archaeologists.

When establishing the first **linear wear - frequency model**, our fundamental starting points are the **Archard equation**, which describes the relationship between wear amount and applied force, and the study of the **force distribution on the sole of the foot** during ascending and descending the stairs. The core concept is to **reproduce** the data of the real wear curve through multiple simulated footsteps on a computer. The data - processing method involves meshing the staircase surface and conducting a **finite - element analysis**.

For the first three problems, we transform the problems into finding the number and position distribution of stepping points that minimize the difference from the real values. For this NP - hard problem, we employ a **genetic algorithm** to solve it. By analyzing information such as the distribution of stepping points in the optimal solution, we can address the first three problems.

For problems 4 to 6: To better describe the impacts of various complex factors on the staircase, we introduce a time variable based on the linear wear - frequency model from the first three problems and establish an **accelerated wear - frequency model**. **For problem 4**: Given the known distribution, if the error between the theoretical solution and the real data is small, the information is considered valid. **For problem 5**: We repeatedly calculate the optimal solution that best fits the theory and reality and use hypothesis testing for judgment. **For problem 6**: Regarding whether the staircase has been renovated, we propose two methods: the gradient method and the regional method. To estimate the renovation time point, we can make a judgment based on **the positive correlation between the renovation time of the renovated area and the wear**.

For problem 7: When determining the source of materials, we encounter a large amount of information. In response, we propose a **similarity evaluation model based on the entropy - weight method**. After **quantifying** each evaluation index, we introduce **entropy weights** and finally define the similarity using the **reciprocal of the Euclidean distance** to evaluate reliable speculations.

For problem 8: In this problem, we no longer use the genetic algorithm for solution. Instead, we simply propose using **the ratio of the standard deviation of the wear position distribution to the average wear depth** as a measure of the dispersion degree of stepping points, and judge the usage pattern of the staircase based on the magnitude of this measure in the actual data.

Finally, we verify the reliability of all the above models, evaluate the advantages, disadvantages, and optimization directions of our models, and write a complete set of solutions in a letter to archaeologists to assist them in inferring information related to staircases.

Keywords: Archaeology, Staircase, Wear - Frequency Model, Genetic Algorithm, Finite - Element Method

Contents

1 Introduction

1.1 Problem Background

Natural materials such as marble, basalt and cedar wood are widely used in civil engineering due to their desirable physical and mechanical properties, including strength, texture, and color. These materials are commonly used on staircases, flooring, and wall coverings.Mohammed et al., 2021

Figure 1: worn stone steps

Staircases are integral components of buildings, and walking habits, as well as usage frequencies, vary across structures. Despite the hardness and durability of these materials, prolonged human activity during the lifetime of a building leads to uneven and varying degrees of wear. These wear patterns provide archaeologists with valuable insights into estimating the construction period of staircases and buildings.

1.2 Restatement of the Problems

Considering the background information and the restricted conditions identified in the problem statement, our group will accomplish the following tasks:

- Predict the frequency of stairs usage.
- Analyze the preferred walking directions of individuals using the stairs.
- Estimate the number of people using the stairs simultaneously.

Using the existing estimate of the building's age, an understanding of stair usage patterns, and general insights into daily activities within the building, provide guidance to address the following questions.

- Analyze the discrepancies between actual wear patterns and known data.
- Predict the age of the stairs and assess the reliability of the estimate.
- Identify any maintenance or renovation work conducted on the stairs.
- Determine the origin of the materials used in the stairs.
- Evaluate human traffic patterns by analyzing the number of people using the stairs throughout the day.

1.3 Our Work

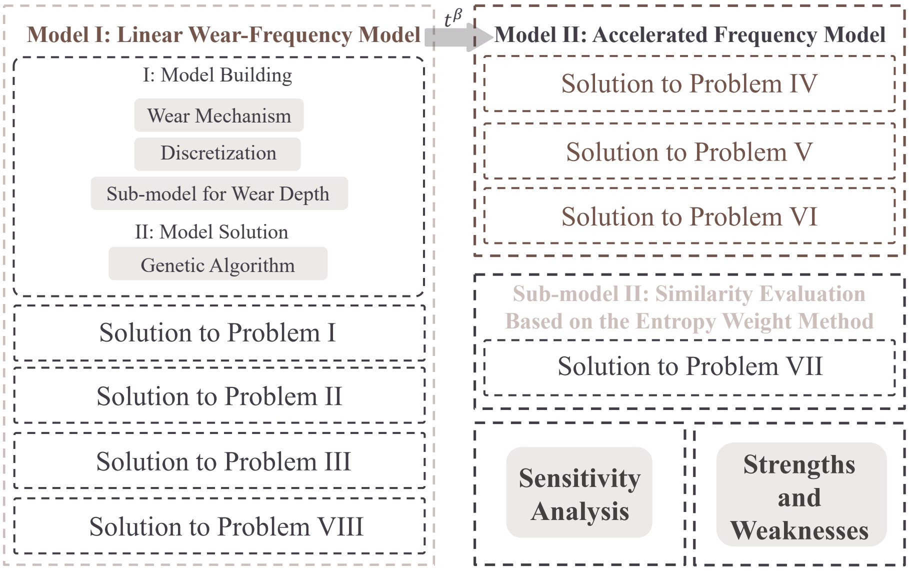


Figure 2: Overview of this work

2 Assumptions and Justification

To simplify the problem and make it convenient for us to simulate real-life conditions, we make the following basic assumptions, each of which is properly justified.

- **Assumption 1:** The wear per step is positively correlated with the normal load in a linear fashion.

  Justification: Marble was selected as the step material, and the wear follows the Archard equation, in which the normal load is linearly related to wear. Therefore, we assume that the wear per step is positively correlated with the normal load in a linear fashion.

- **Assumption 2:** Cumulative wear is positively correlated with the number of steps.

  Justification: Although environmental factors, such as weathering and rainwater erosion, can contribute to step wear, their effect is minimal compared to the impact of stepping. Therefore, we primarily consider wear caused by stepping, where cumulative wear is positively correlated with the number of steps.

- **Assumption 3:** Both feet exert equal force while walking, and the probability of stepping with the left or right foot is statistically identical.

  Justification: Assuming the individual is in good health, both the left and right feet exert equal force while walking. In the specific context of ascending and descending stairs, we assume that the probability of stepping with either foot is the same.

- **Assumption 4:** When a staircase is first constructed and has not undergone any renovation, it can be considered homogeneous.

  Justification: In most cases, people tend to use building materials with consistent textures, as this allows designers to perform structural analysis and calculations more easily without accounting for internal material heterogeneity. Additionally, it helps to avoid complex stress distributions and deformation issues caused by material inconsistency.

- **Assumption 5:** The wear caused by each individual during ascent is identical, though wear during ascent and descent differs. of solving the problem.

  Justification: In our analysis, we will use average body weight, average foot size, and average sliding distance using, as the large number of stair users typically results in no statistically significant differences. Due to variations in muscle activity during ascent and descent, the normal loads exerted on the stairs differ, leading to distinct wear patterns for ascending and descending movements.

3 Notations

Table 1: the list of notation

Symbol	Meaning
V	total wear of the step
$d(x, y)$	the wear depth
$f(x, y)$	the footfall frequency intensity
γ	the wear caused by a single step
α	a matrix that represents the effect to the whole stair

where we define the main parameters while specific value of those parameters will be given later.

4 Model I:Linear Wear-Frequency Model

4.1 Model Building

4.1.1 Wear Mechanism

When two surfaces come into contact, even if there is no movement, there are molecular interactions. These interactions create bonds that are broken when there is movement of one surface across or away from the other surface. This is adhesive wear. The movement and subsequent breaking of bonds releases energy in the form of frictional heat and loosens material from one or both surfaces.

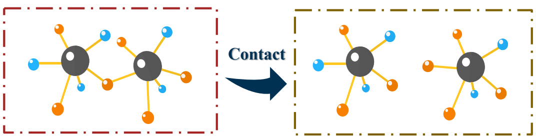


Figure 3: Wear Mechanism

These particles can remain loose between the surfaces, or can become attached to the opposite surface or at another location on the original surface. In the early stages of adhesive wear the damage may not be visible except at very high power magnification.Long-term use will lead to significant volume loss of the step.

4.1.2 Discretization of the Step's Surface

Assuming the step is an ideal rectangular prism when initially used, and stepping occurs on the upper surface, the sole of the shoe comes into contact with the surface, causing wear. A rectangular coordinate system is established on the upper surface of the step, with the longer side as the x-axis and the shorter side as the y-axis. The upper surface is divided into small squares, as shown in the figure below.

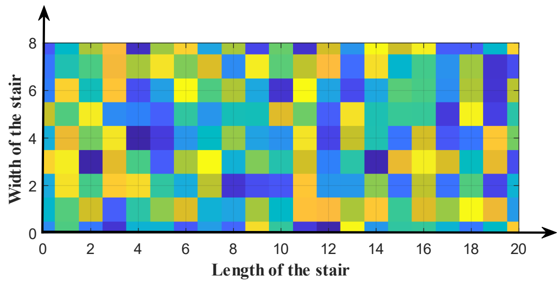


Figure 4: Diagram of Step Discretization

Let (i, j) represent the index of a square, where $i \leqslant M$ and $j \leqslant N$, with M and N being the maximum indices along the length and width, respectively. For example, $(1, 1)$ corresponds to the square in the first row and first column.

Bergstra's method is used to divide the foot into regions and extract the peak pressure and pressure-time integral (PTI) for each region.Cho et al., 2021

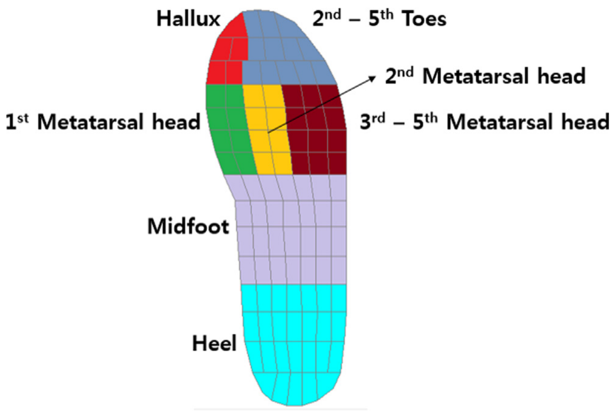


Figure 5: The mask composed of seven regions.

4.1.3 Sub-model for Wear Depth

To model wear, a sub-model for wear depth is constructed. Given the discretization of the step, wear in each square can be represented by the wear depth $d(x, y)$, where x and y are the indices of the square. Therefore, the total wear of the step can be expressed as

$$V = \sum_{i=1}^{M} \sum_{j=1}^{n} d_{i,j} \Delta S = \gamma_1 U_1 + \gamma_2 U_2 \tag{1}$$

where V is the total volume of wear debris produced, U_1 is the total number of upward steps, U_2 is the total number of downward steps, γ_1 is the wear caused by a single upward step, and γ_2 is the wear caused by a single downward step.

According to the Archard equationKaraca et al., 2012

$$V_{\text{Single}} = \frac{kFs}{H} \tag{2}$$

where k is a dimensionless constant, F is the total normal load, s is the sliding distance, H is the hardness of the softest contacting surfaces.

Since the sliding distance is much smaller than the grid step size, the wear V within the grid (i, j) depends primarily on the force F. Thus, the wear depth caused by a single upward step and a single downward step in the grid (i, j) can be expressed as

$$d_{1,in(i,j)} = \frac{k}{H\Delta S}\left(F_{1,in(i.j)} \cdot s_1\right) \tag{3}$$

$$d_{2,in(i,j)} = \frac{k}{H\Delta S}\left(F_{2,in(i.j)} \cdot s_2\right) \tag{4}$$

A single step may cause wear on more than one grid. he impact of a single step on the total wear depth of the entire step is represented by $\boldsymbol{\alpha}_{1L}(x, y)$, $\boldsymbol{\alpha}_{1R}(x, y)$, $\boldsymbol{\alpha}_{2L}(x, y)$, $\boldsymbol{\alpha}_{2R}(x, y)$.

d can be represented by

$$d = \sum_{i=1}^{m}\sum_{j=1}^{n}\left[\boldsymbol{\alpha}_{1L}(i,j)\boldsymbol{U}_{i,j}^{1L} + \boldsymbol{\alpha}_{1R}(i,j)\boldsymbol{U}_{i,j}^{1R} + \boldsymbol{\alpha}_{2L}(i,j)\boldsymbol{U}_{i,j}^{2L} + \boldsymbol{\alpha}_{2R}(i,j)\boldsymbol{U}_{i,j}^{2R}\right] \tag{5}$$

where $\boldsymbol{U}_{i,j}$ represents The number of times a square is the center square.

Let $\boldsymbol{U}_{i,j} = [u_{i,j}]$,

$$U_{1L} = \sum_{i=1}^{m}\sum_{j=1}^{n} u_{i,j}^{1L}, U_{1R} = \sum_{i=1}^{m}\sum_{j=1}^{n} u_{i,j}^{1R}, U_{2L} = \sum_{i=1}^{m}\sum_{j=1}^{n} u_{i,j}^{2L}, U_{2R} = \sum_{i=1}^{m}\sum_{j=1}^{n} u_{i,j}^{2R} \tag{6}$$

U_1 and U_2 can be represent by

$$U_1 = U_{1L} + U_{1R}, U_2 = U_{2L} + U_{2R} \tag{7}$$

From Assumption 3,

$$U_{1L} = U_{1R}, U_{2L} = U_{2R} \tag{8}$$

The probability distribution matrix can be derived as

$$\mathbf{F}_{1L} = \frac{1}{U_{1L}}\mathbf{U}^{1L}, \mathbf{F}_{1R} = \frac{1}{U_{1R}}\mathbf{U}^{1R}, \mathbf{F}_{2L} = \frac{1}{U_{2L}}\mathbf{U}^{2L}, \mathbf{F}_{2R} = \frac{1}{U_{2R}}\mathbf{U}^{2R}. \tag{9}$$

4.2 Model Solution

During the modeling process, many parameters to be determined were generated, and in order to determine the specific values of these parameters, we review a large number of literature and online resources.

The evaluation criteria are based on four dimensions: the x-coordinate, the y-coordinate, the direction (upward or downward), and the foot (left or right). Due to the large number of x and y values, the solution space is highly complex, and no prior information is available. Therefore, a **genetic algorithm** is employed to solve the problem, as it is well-suited for handling the large-scale and complex nature of the model.

According to Assumption 1 and 2, on the stair surface xOy, the wear depth$d(x, y)$(in meters) at each point and the footfall frequency intensity $f(x, y)$(in m^2)are related as follows:

$$d(x, y) = k_1 f(x, y) + \epsilon(x, y) \tag{10}$$

where k_1 can be calculated by 2, $\epsilon(x, y)$ accounts for the wear depth error caused by non-footfall factors, such as the natural environment.

According to Assumption 3, the objective is to minimize the mean square error of the wear depth. For computational convenience, the stair surface is discretized into $p \times q$ square grids, each with a side length l, where all grids have the same value. That is, x and y are discretized as non-negative integers. The problem then becomes the task of solving for:

$$\begin{aligned} \min \quad & \sum_{x,y} (d(x, y) - k \cdot f(x, y))^2 \\ \text{s.t.} \quad & \begin{cases} 0 \leq x \leq p - 1 \\ 0 \leq y \leq q - 1 \end{cases} \end{aligned} \tag{11}$$

To solve this equation, a genetic algorithm is used:

- According to the British Standard (BS 5395-1)Vesela, 2019, the step width must lie within the range of $30cm \leqslant L_2 \leqslant 45cm$. Thus, L_2 is set to 40 cm, and L_1 is chosen as 150 cm.
- Given that the foot length is approximately 30 cm and the width is about 10 cm, which is divided into a 24 × 8 grid, the step length is selected as 1.25 cm, resulting in a 120 × 32 grid.
- Marble is chosen as the material for the steps, with a Mohs hardness of 2.0–4.3. Based on relevant data, the hardness is taken as H = 2 GPa, and the wear coefficient (k) is set to 160. The calculated results are compared with the sample M_1 from the literature [6], and the comparison confirms their validity. The sliding distance is set to 3 mm.

The figure below presents the discretization results of the foot sole, illustrating a heatmap of peak pressure during both ascending and descending stairs. The contact wear is linearly related to the normal pressure, and the peak pressure matrix serves as the foundation for the subsequent wear calculation.

Table 2: Genetic Algorithm for Problem I

Step	Operation
1	Set the population size M, the maximum number of generations G, the crossover probability p_c, and the mutation probability p_m.
2	Based on the problem setup, each feasible solution is a $p \times q$ matrix. Randomly initialize the values of M feasible solutions.
3	Compute the mean square error for the M feasible solutions.
4	Select the two solutions with the smallest mean square error as parents.
5	Perform single - point crossover by randomly selecting a point in the matrix and exchanging values at the corresponding positions in the parent matrices.
6	Apply mutation with a certain probability by randomly selecting points and resetting their values.
7	Recalculate the mean square error for both the parents and the offspring. If the desired result is achieved or the maximum number of generations is reached, proceed to step 8; otherwise, return to step 4.
8	After obtaining $f(x,y)$, compute the total frequency $f = \sum_{x,y} f(x,y)$

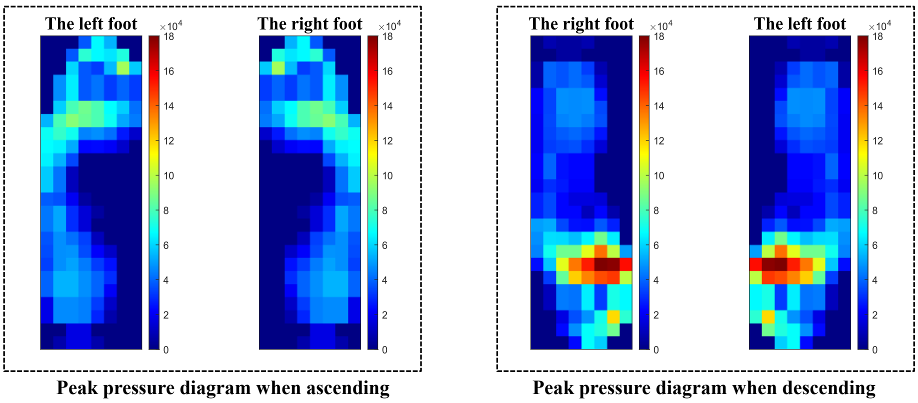


Figure 6: Peak pressure simulation when ascending and descending

The total pressure applied by the foot was standardized, assuming that if the foot does not fully contact the step, the total pressure remains the same as when the foot is entirely in contact with the step.He et al., 2023

4.3 Solution to Problem I

4.3.1 Constraint

According to 1 and 8, the constraint can be represented as

$$2\gamma_1 * U1 + 2\gamma_2 * U2 = \sum_{x=1}^{m} \sum_{y=1}^{n} \mathbf{d}_{x,y} \tag{12}$$

When the ascending-to-descending ratio is fixed, that is

$$\text{ratio } = \frac{U_1}{U_2} \tag{13}$$

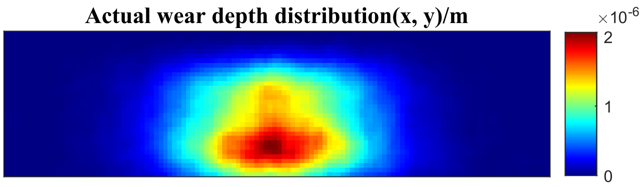
The equation becomes a system of first order equations with one variable, thus the frequency can be determined by this constraint. In our subsequent simulations, we measured and calculated the total wear volume, which enabled us to derive the specific constraint.

4.3.2 Model Application

According to the system block diagram of Model I, we can infer that the wear depth distribution of the stairs can be obtained through measurement. By measuring and referencing relevant data, the material of the stairs can also be determined. A genetic algorithm is employed to solve for the probability distributions of ascending and descending footstep points.

Marble stairs are selected as the analysis case, with specific parameters provided in the model solution. The wear depth distribution of the stairs is assumed to follow a Gaussian distribution based on position i, with added noise. It is also assumed that the probabilities of ascending and descending are equal. Over the course of one month, 6,000 people step on the stairs.

For real measurement data, we use a genetic algorithm, while for predicted data, we assume the wear data follows a Gaussian distribution. Based on this assumption, we solve for the wear depth heatmaps in both cases.Fujiyama and Tyler, 2010



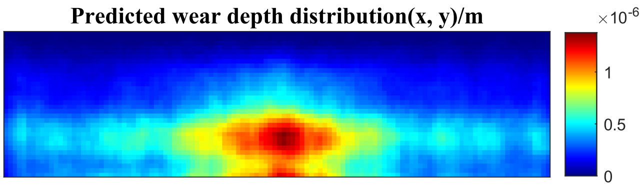


Figure 7: Wear depth distribution

From the wear depth data, we can infer the footstep distribution. The heatmaps of the actual and predicted footstep center distributions are shown below.

In the assumptions of Problem I, it is assumed that the probabilities of ascending and descending are equal. Thus, based on the actual and predicted footstep center distributions, we obtain the distributions of the actual and predicted ascending and descending footstep centers. The corresponding heatmaps are shown below.

One key dimension of the feasible solution is the ascending/descending decision. By analyzing the ascending/descending ratio, we found a predicted ratio of 0.9953 and an actual ratio of 0.9698, with an error of 2.6%. This confirms the feasibility of the method.

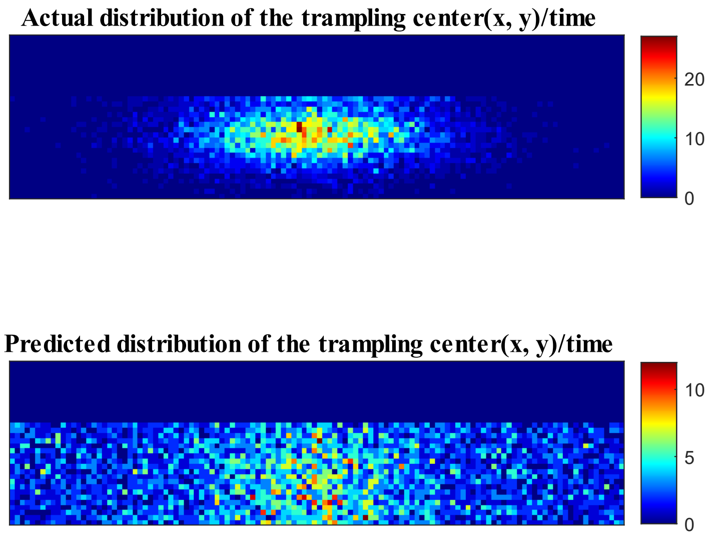


Figure 8: Distribution of the trampling center

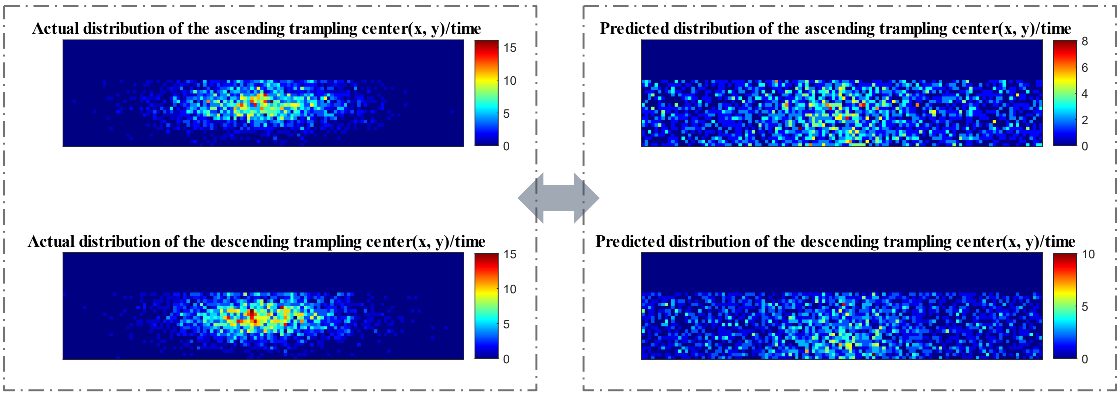


Figure 9: Distribution of the ascending and descending trampling center

4.3.3 Frequency Calculation

Assuming a 30-day time window, in our data setup, the actual frequency of stair usage is calculated as follows.

$$frequency = \frac{f_N}{T} \tag{14}$$

where f_N represents the time people step on the stairs, which has a specific value of 6000 and T represents the time window we set, which is 30 days. Substituting the specific values, the final result is 200.

The predicted frequency can be calculated as follows. According to 13, we have

$$\begin{cases} U_1 = \frac{V}{2\left(\gamma_1 + \frac{\gamma_2}{ratio}\right)} \\ U_2 = U_1 \times ratio \end{cases} \tag{15}$$

then the frequency can be represented by

$$frequency = \frac{2\left(U_1 + U_2\right)}{T} \tag{16}$$

The result of the calculation is 199.6384, with an error of 0.18% between the two values, indicating a good match. This demonstrates that the proposed solution effectively addresses the problem.

Table 3: Solution of Problem I

Frequency Type	**Value of Frequency**
Actual Frequency	200
Predicted Frequency	199.6384

4.4 Solution to Problem II

4.4.1 Model Application

According to Hypothesis 5, assume there are m reference points for ascending, which implies there are $f - m$ reference points for descending, leading to the following relationship:

$$\begin{aligned} f(x,y) =& k_2\left(\beta\left(\sum_{h=0}^{\frac{m}{2}-1}\alpha_{ur}(x_h,y_h) + \sum_{i=0}^{\frac{m}{2}-1}\alpha_{ul}(x_i,y_i)\right)\right. \\ &\left. + \sum_{j=0}^{\frac{f-m}{2}-1}\alpha_{dr}(x_j,y_j) + \sum_{k=0}^{\frac{f-m}{2}-1}\alpha_{dl}(x_k,y_k)\right) \\ &+ \varepsilon(x,y) \end{aligned} \tag{17}$$

thus, similar to the first problem, the task is to be

$$\begin{aligned} \min \quad & \sum_{x,y}\left(f(x,y) - k_2\left(\beta\left(\sum_{h=0}^{\frac{m}{2}-1}\alpha_{ur}\left(x_h,y_h\right)\right.\right.\right. \\ & \left.\left.\left. + \sum_{i=0}^{\frac{m}{2}-1}\alpha_{ul}\left(x_i,y_i\right)\right) + \sum_{j=0}^{\frac{f-m}{2}-1}\alpha_{dr}\left(x_j,y_j\right) + \sum_{k=0}^{\frac{f-m}{2}-1}\alpha_{dl}\left(x_k,y_k\right)\right)\right)^2 \\ \text{s.t.} \quad & \begin{cases} \frac{3a}{2} \le x_h, x_k \le p - \frac{a}{2} \\ \frac{a}{2} \le x_i, x_j \le p - \frac{3a}{2} \\ 0 \le y_h, y_i, y_j, y_k \le ql - \frac{b}{2} \end{cases} \end{aligned} \tag{18}$$

the task is to determine m, where β represents the correction factor introduced by the unequal pressures of ascending and descending.

As in Problem 1, a genetic algorithm is used to solve this. The only difference from Problem 1 is the data structure. The feasible solutions in this case are represented by a $f \times 1$ vector, where each element is in the form (x_i, y_i, m), Here, m is a binary decision variable: when $m = 1$, it indicates ascending, and the value is taken from $\alpha_1(x, y)$;when $m = 0$, it indicates ascending, and the value is taken from $\alpha_2(x, y)$.

4.4.2 Direction Favor

The pressure on the stairs varies between ascending and descending. As a result, different ascending-to-descending ratios lead to variations in stair wear. Three distinct scenarios have been defined, with the specific ratios presented in the table with the total number of people remaining the same.

Table 4: Value adoption in ascending and descending probabilities

Number of condition	Ratio
Condition 1	1:1
Condition 2	1:3
Condition 3	3:1

The heatmaps below show the distribution of wear depth and trampling centers for Condition 2 and Condition 2 as Condition 1 has been illustrated in last subsection.

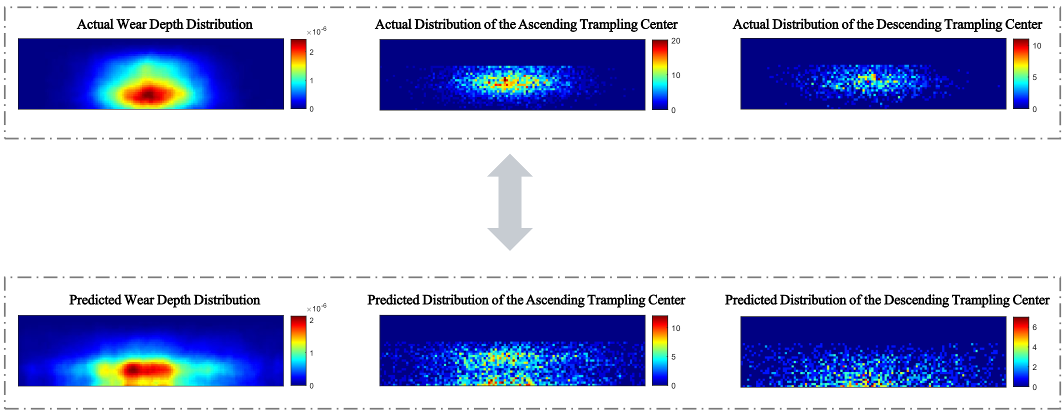


Figure 10: Heatmap when ratio=1:3

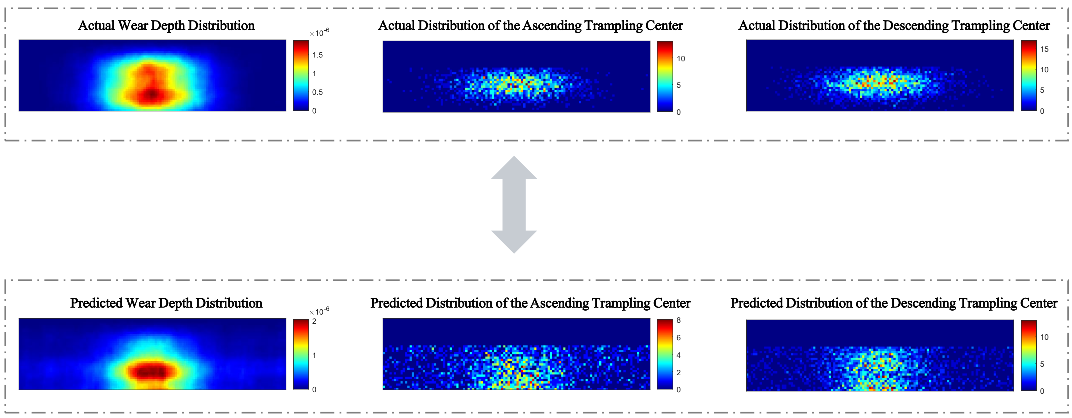


Figure 11: Heatmap when ratio=3:1

The previous section highlighted that the pressure exerted on the stairs differs between ascending and descending traffic, and varying ratios of these directions result in different levels of wear. We simulated the variation in total wear as a function of the proportion of descending traffic. Under a constant pedestrian flow, the preference for ascending or descending is directly correlated with the total wear.

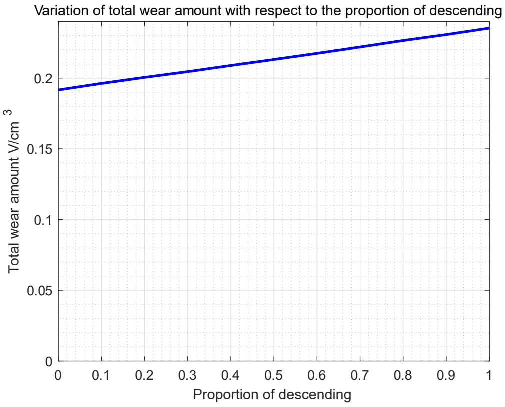


Figure 12: Variation of total wear amount with respect to the proportion of descending

Consequently, the corresponding results can be derived from this curve. For instance, with a fixed total pedestrian flow, higher wear indicates a greater use of the stairs for descending, whereas lower wear suggests more use for ascending.

4.5 Solution to Problem III

We assume that only pairs of people walking side by side are considered in this scenario. If the length of a stair step falls within the range [40, 80], it is classified as single walking; if it falls outside this range, it is considered parallel walking, where both feet exert pressure on the step with each stride. The pedestrian flow is set to 6000, and the predicted heatmap is generated. The actual heatmap is obtained using a genetic algorithm.

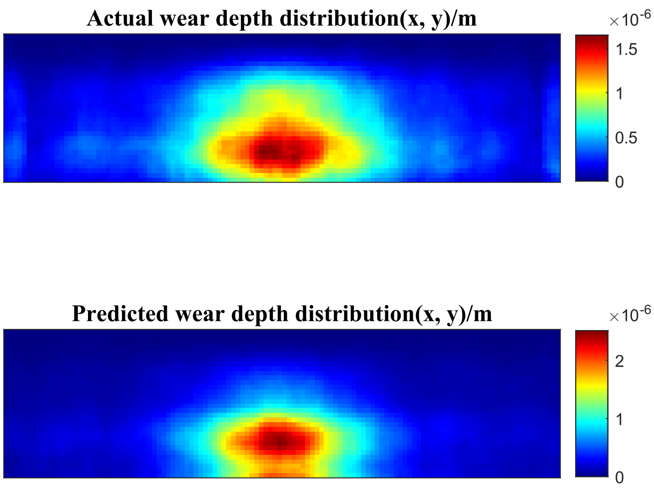


Figure 13: Wear depth distribution

The predicted ratio of ascending to descending traffic is 0.6295, while the ratio derived from

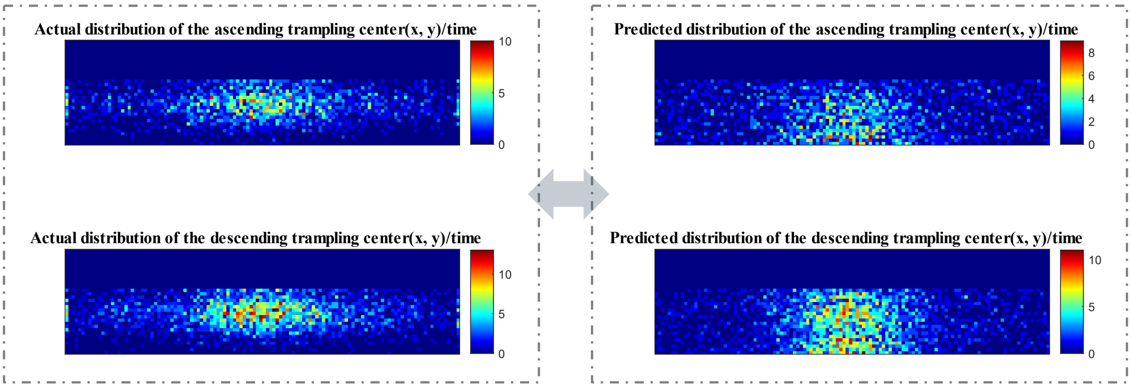


Figure 14: Distribution of the ascending and descending trampling center

the genetic algorithm is 0.6081, yielding a 3.5% error, which confirms the method's feasibility. The parallel walking ratio is defined as the number of times stepping on the edge divided by the number of times stepping in the middle in this model as 23.8%.

Since we assume that both feet exert pressure on the steps during parallel walking, the ratio of parallel walkers to solo walkers is approximately 1:8, indicating a **preference for walking alone**.

4.6 Solution to Problem VIII

4.6.1 Assumptions and Model Setup

Consider a staircase with m steps, where each step can accommodate up to n people side by side . The instantaneous pedestrian flow on the staircase is classified as follows:

- If the instantaneous pedestrian flow is less than m, it is considered significantly low frequency.
- If the instantaneous pedestrian flow exceeds $\frac{mn}{2}$, it is considered significantly high frequency.
- Pedestrian flow between these two extremes is considered normal.

4.6.2 Model Establishment

Study of Daily Pedestrian Flow: according to 1,

$$N = \frac{\sum_{x,y} d(x,y)}{n \sum_{x,y} \beta\alpha(x,y)} \tag{19}$$

As pedestrian flow increases, more people will stand on each step simultaneously. Consequently, the standard deviation of wear on the staircase, which is typically associated with high pedestrian flow over short periods, will increase. Conversely, the standard deviation will decrease under lower pedestrian flow. To measure the usage pattern with standard deviation, we define $C = \frac{\sigma}{d_{\text{avg}}}$ as the dispersion of footfalls, where d_{avg} represents the average depth of footfalls on the staircase surface, depending solely on the number of footfalls. d_0 also represents the average depth of a single footfall. Greater dispersion indicates more scattered footfalls, which can be used to distinguish between usage patterns characterized by short bursts of high flow or long periods of low flow.

Analysis of the Boundary Value of C: Under the assumption that the instantaneous pedestrian flow is m, it can be considered that there is only one person on each step. Since each step can accommodate a maximum of n people, and the horizontal space occupied by each person is $\frac{l}{n}$ (where l is the step length), we assume that all individuals are standing in the center of the step. According to 3σ principle, $3\sigma = \frac{l}{2n}$. Under this assumption, at an instantaneous flow of $\frac{mn}{2}$, each step accommodates $\frac{n}{2}$ people, with the total horizontal space occupied being $\frac{l}{2}$, which corresponds to individuals standing in the center of the step and $3\sigma = \frac{l}{4}$.

When $C < \frac{l}{fd_0} \cdot \frac{1}{6n}$, the staircase is primarily used in a significantly low-frequency mode with long-term, low pedestrian flow. When $C > \frac{l}{fd_0} \cdot \frac{1}{12}$, the staircase is primarily used in a significantly high-frequency mode with short-term, high pedestrian flow. Pedestrian flow between these two extremes is considered neither significantly high nor significantly low.

4.6.3 Model Application

After performing the calculations, the **daily usage frequency** is found to be 29.9946, and the **footfall dispersion** is 594.8949. A footfall dispersion below 240 is considered significantly low pedestrian flow, while a dispersion above 480 is considered significantly high. With a standard deviation of $\frac{mn}{2}$, the pedestrian flow is categorized as significantly high.

5 Model II:Accelerated Wear-Frequency Model

5.1 Model Building

We define a one-dimensional time coordinate system with the present as the origin, extending backward in time. Here, t represents the number of time units from the present. We introduce a new function,

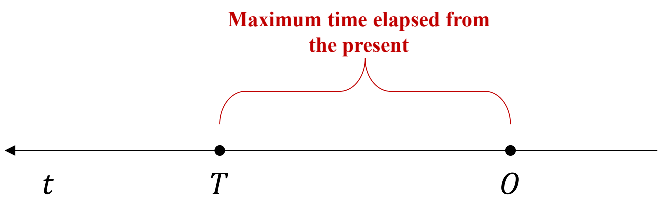


Figure 15: The time coordinate

$$\boldsymbol{\alpha}_{1L}(x, y, t) = \boldsymbol{\alpha}_{1L}(x, y) \cdot t^{\beta} \tag{20}$$

where β is a constant, experimentally determined, and set to 0.1, reflecting the material's variation characteristics. The input to α consists of three parameters, and the output is a two-dimensional matrix.

$$\vec{d} = \sum_{\tau=1}^{T}\sum_{i=1}^{m}\sum_{j=1}^{n}\left[\vec{\alpha}_{1L}(i,j,\tau)U_{i,j,\tau}^{1L} + \vec{\alpha}_{1R}(i,j,\tau)U_{i,j,\tau}^{1R} + \vec{\alpha}_{2L}(i,j,\tau)U_{i,j\tau}^{2L} + \vec{\alpha}_{2R}(i,j,\tau)U_{i,j,\tau}^{2R}\right], \tag{21}$$

$\mathbf{U}^{1L}$ is then extended into a three-dimensional matrix, where $U_{i,j,\tau}^{1L}$ represents the number of times a footfall center occurs at grid point,(i,j) over the past τ time units. T represents the maximum time elapsed from the present at any given stair step.

5.2 Solution to Problem IV

Some information about the building has already been collected, including an estimate of its age, the usage patterns of the stairwell, and an estimate of the daily life patterns within the structure. We assume the building is 120 months old. Based on the stairwell usage and daily life patterns, a pedestrian flow matrix Q can be derived. For simplicity, we assume that Q follows a Gaussian distribution and, for generality, that the ratio of ascending to descending pedestrians is 1:1. In Model II, the gradual material changes due to long-term use are accounted for, with these non-linear changes represented by t^{β}.

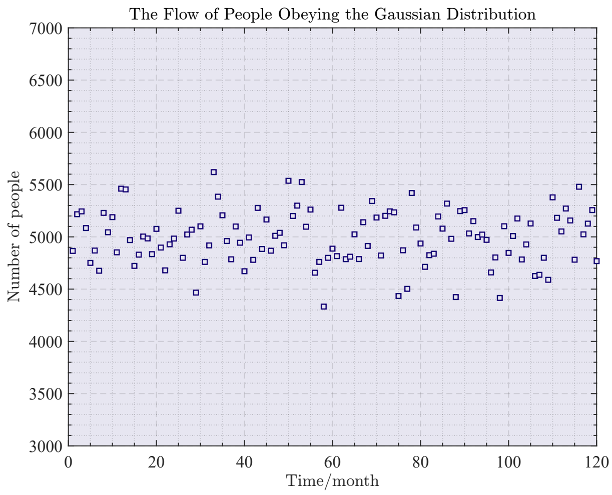


Figure 16: Flow of people obeying the Gaussian Distribution

In our program, we determine the probability distribution matrix for the footfall centers (with decimal values possible for matrix elements). By multiplying the pedestrian flow matrix by the probability distribution matrix and inputting the result into the model, we obtain the wear depth distribution. All probability distributions are Gaussian, with the assumption that ascending pedestrians are on the right side and descending pedestrians on the left (with a consistent center along the y-axis).

Through simulation, we generate both the theoretical and actual wear depth heatmaps.

By subtracting the two matrices, we obtain the error distribution matrix, which is then optimized. The error distribution matrix is normalized, and a threshold is introduced to measure the consistency of the information. If the normalized error E_{norm} is below the threshold, it is considered consistent; otherwise, it is deemed inconsistent.

The normalized error is closely related to the average flow Q_{avg}. By comparing the curves of the uniform and Gaussian distributions, a logarithmic relationship is identified and fitted to determine the appropriate threshold.

Through linear regression and first-order polynomial fitting, the fitting result is obtained as

$$E_{\text{norm}} = -0.0429 \ln \left(Q_{\text{avg}} \right) + 0.4249 \tag{22}$$

By incorporating a margin based on the fitting, a reasonable threshold can be determined.

5.3 Solution to Problem V

5.3.1 Time Estimation

The known usage time is denoted as T, and we estimate the usage time based on it. A predicted random variable for the stair usage time, t, is introduced. The optimization model process is outlined as follows:

From T, the pedestrian flow for ascending and descending, Q_1 and Q_2, can be derived. The random variable t also corresponds to $Q_1(t)$ and $Q_2(t)$, both of which follow a Gaussian distribution. The corresponding wear depth matrices, d_{theo} and d_{act}, are obtained, and the error distribution function is derived by subtracting the two. The optimization goal is to minimize this error.

$$\vec{e} = \left| \overrightarrow{d_{\text{act}}} - \overrightarrow{d_{\text{theo}}} \right| \tag{23}$$

$$\min e'_{\text{norm}} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} e_{\text{norm},i,j}}{\sum_{i=1}^{n} \sum_{j=1}^{m} d_{\text{act},i,j}}, t \in N^{+} \tag{24}$$

The optimal solution t_{opt} and the minimum normalized error E_{opt} are then analyzed in relation to the variation in the average pedestrian flow.

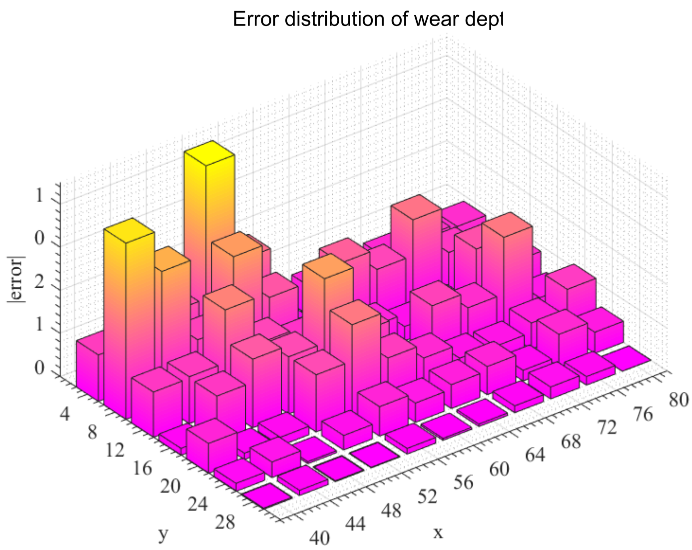


Figure 17: Error distribution of wear depth

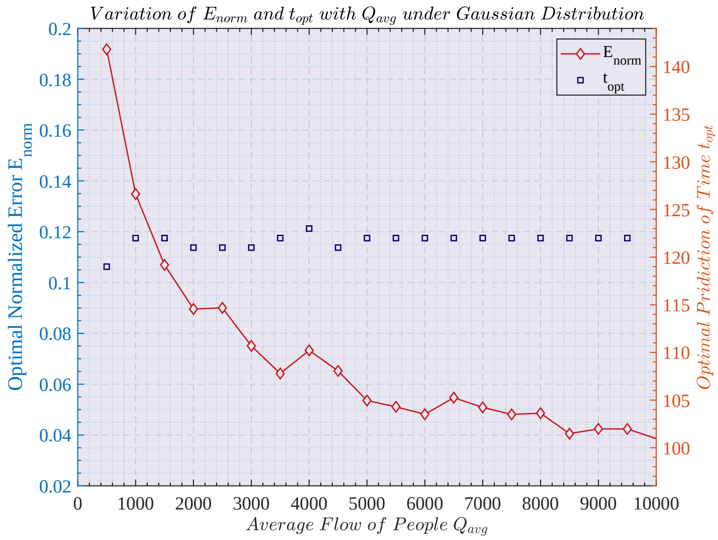


Figure 18: Variation under Gaussian Distribution

It can be observed that as the pedestrian flow increases, the minimum error gradually decreases, and the optimal time solution becomes smoother and slightly higher than the actual time.

5.3.2 Reliability Evaluation

Since t is a random variable influenced by many factors—such as inaccuracies in material properties, errors in the average pedestrian flow per unit time, and discrepancies in the footfall center probability distribution—and these factors are not the primary influencing elements, it is assumed that t approximately follows a Gaussian distribution. Therefore, when the pedestrian flow is assumed to follow a Gaussian distribution, the average pedestrian flow, material characteristics, and footfall center probability distribution are varied. The optimal solutions obtained are recorded as observations.

The confidence interval for T at a $99\%(\alpha = 0.01)$ confidence level is given as $\left[\bar{X} \pm \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)\right]$, where $\bar{X}$ is the mean of the observed values t_{ob}, n is the number of observations, S is the sample standard deviation $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}$, and $t_{\frac{\alpha}{2}}(n-1)$ is the t-distribution critical value at the $1-\alpha$ quantile with $n-1$ degrees of freedom (i.e., the upper region of the t-distribution corresponding to $\alpha/2$).

Using this condition, the minimum estimated age T_{min} is 118.2666, and the maximum estimated age T_{max} is 121.6668. The probability that the actual time T falls outside this interval is only 1%.

5.4 Solution to Problem VI

The footfall probability distribution can be determined by assuming an equal ratio of ascending to descending traffic, while only considering single-use scenarios. To simplify the analysis, we assume that only one renovation has occurred in the area. The problem is then restated as identifying the "renovated area" and the "renovation time point" (i.e., how long ago the renovation took place). Assuming the pedestrian flow Q follows a normal distribution, a more accurate result for the stair usage time T_{original} can be obtained directly using the method described in Problem V.

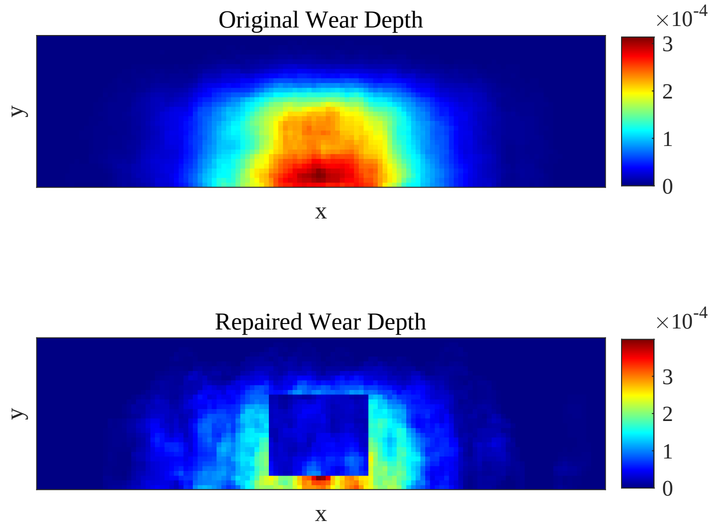


Figure 19: Comparison between worn and repaired stair

We propose two solutions——**Gradient Method** and **Region Method** to analyze the renovation detection requirements.

5.4.1 Gradient Method

The principle behind the gradient method is that the gradient at the interface between new and old materials is relatively large, which can be used to describe the boundary. This method does not require prior knowledge of certain conditions. We assume that renovation resets the wear depth to zero. The gradient maps for both the unrenovated and renovated steps are shown below.

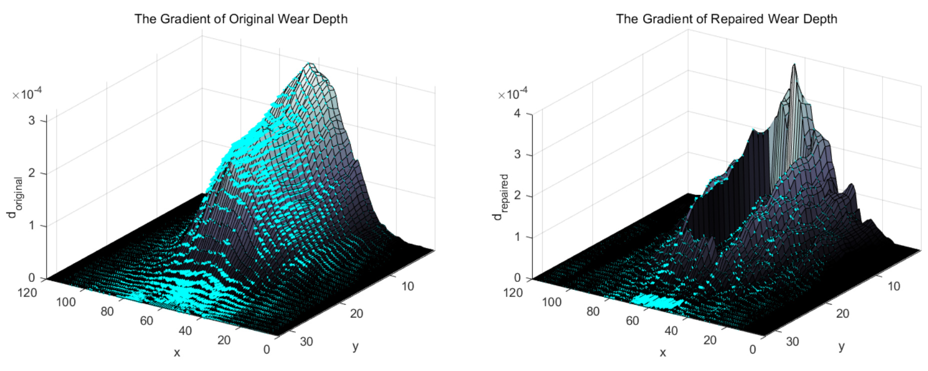


Figure 20: Gradient of worn depth

The renovated steps show a significant flat area at the center, generating a large gradient.

Evaluation: The advantage of the gradient method lies in its simplicity and convenience, as it does not require model calculations. However, its main drawback is that it becomes less accurate, especially when the a_factor is large or the renovation time $t_repaired$ is long. In these cases, predictions become highly inaccurate. Additionally, the method's accuracy decreases in areas where the original wear depth was minimal.

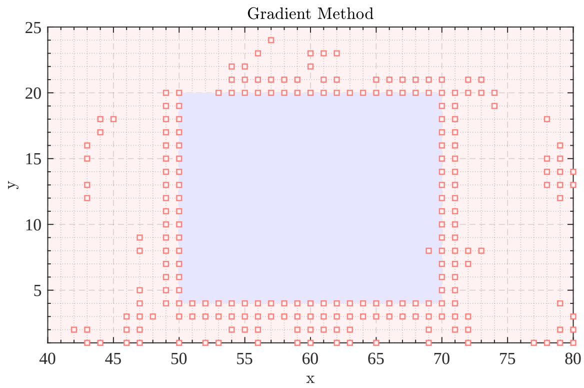


Figure 21: Boundary detection diagram

5.4.2 Region Method

Based on this, we propose the Region Method for analysis.

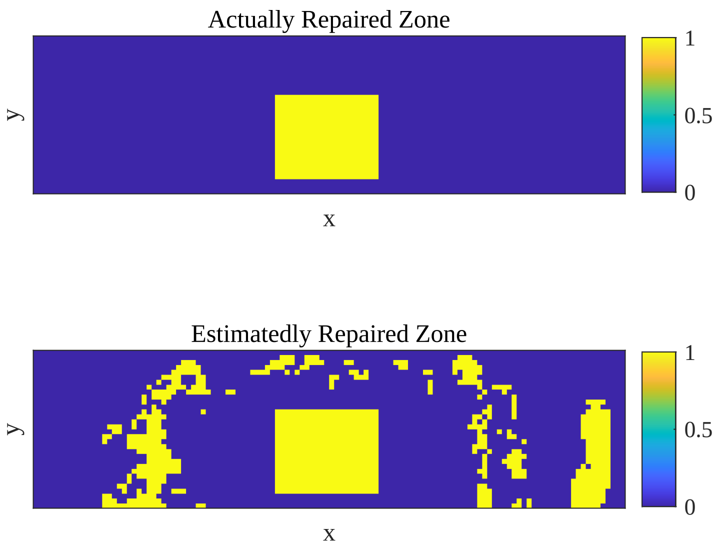


Figure 22: Repaired zone

Evaluation: The advantage of the region method is its higher prediction accuracy in specific regions, as it is less prone to boundary misclassification. It also performs well in situations where the renovation time is extensive or the material properties exhibit significant differences. However, it requires additional information (e.g., the probabilities of ascending and descending traffic, and the footfall center probability distribution). If this information is inaccurate, it will negatively impact the predictions.

5.4.3 Estimating Renovation Time

This problem is similar to Problem V and can be solved using the optimization model. Based on 23,24, the reliability analysis graph is shown below:

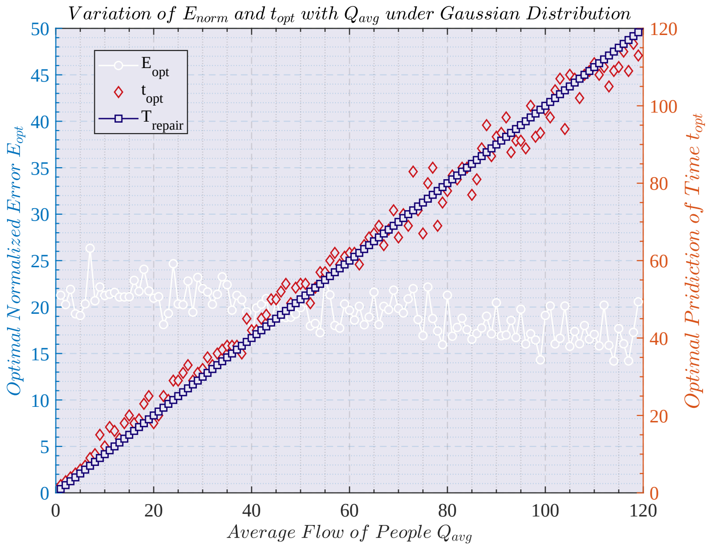


Figure 23: Composite chart

6 Sub-model II: Similarity Evaluation Based on the Entropy Weight Method

Relevant macroscopic physical properties include:Density,Hardness,Color(encoded using color coding),Texture(processed by computer to calculate the mean square error) and Structure.

Relevant geographical information includes: Estimated distance from the material's source to the discovery site (with the distance of the staircase set to 0). Estimated terrain complexity between the material's source and the discovery site (with the terrain complexity of the staircase set to 0).

Relevant historical information includes: Literature records, converted into inferred confidence (with the confidence for the staircase material set to 1).

6.1 Model Building

- Assume that m inferences are made, each with n indicators. For each inference, a feature vector K is constructed. The material used in the staircase is also associated with a feature vector based on the respective indicators.
- To minimize discrepancies in the absolute values of indicators, these $m+1$ feature vectors are synthesized into an $(M+1) \times N$ matrix K. Each column (representing an individual indicator) is then normalized to obtain a new matrix K^{norm}.
- The entropy weight method is applied to assign a weight to each indicator: Calculate the entropy of the j-th indicator as $e_j = -\frac{1}{\ln(m+1)} \sum_{i=1}^{m+1} K_{ij}^{\text{norm}} \ln K_{ij}^{\text{norm}}$. Calculate the coefficient of variation $g_j = 1 - e_j$ for the j-th indicator. Calculate the weight of the j-th indicator, denoted as $w_j = \frac{g_j}{\sum_{j=1}^{m+1} g_j}$.
- The similarity between each sample and the material used in the staircase (represented by K_0^{norm}) is defined as the inverse of the Euclidean distance, with entropy weights applied: $S\left(K_i^{\text{norm}}, K_0^{\text{norm}}\right) = \frac{1}{1+\sqrt{\sum_{j=1}^{n} w_j \left(K_{ij}^{\text{norm}} - K_{0j}^{\text{norm}}\right)^2}}$. The similarity between each sample and

L_0 is then calculated. According to this definition, if the prediction exactly matches L_0, the similarity reaches its maximum value of 1.

- The samples are ranked based on similarity, and the highest similarity is selected. If the similarity exceeds 90% (the threshold defined arbitrarily), the estimation is considered reliable.

6.2 Model Application

As is illustrated, if the similarity exceeds 90% (the threshold defined arbitrarily), the estimation is considered reliable. We assign value randomly, and we get their similarity as 0.9050,0.9290,0.9171,0.8848,0.9048,0.9552,0.9032,0.8932,1.0000. Those who exceeds 90%, the estimation will be considered reliable.

7 Model Analysis and Sensitivity Analysis

Model 1: Linear Wear-Frequency Model **anderson2006**

According to the Archard equation, we have

$$d(x,y) = s\frac{k}{H}\sum_{x,y}\alpha(x-x_i, y-y_i)$$

where k and H are inherent parameters of the material, and s is the surface sliding distance. Errors in these parameters can cause deviations in the model's predictions.

The error is defined as:

$$e = \frac{\sum_{x,y}|d_{\mathrm{cal}}(x,y) - d_{\mathrm{true}}(x,y)|}{\sum_{x,y} d_{\mathrm{true}}(x,y)} = \frac{\sum_{x,y}|s\frac{k}{H}\sum_i \alpha(x-x_i, y-y_i) - d_{\mathrm{true}}(x,y)|}{\sum_{x,y} d_{\mathrm{true}}(x,y)}$$

We only conduct a sensitivity analysis on the measurement errors of the three parameters. When the parameters change slightly, it is verified by the genetic algorithm that the change in the predicted wear distribution $\sum_{x,y}\alpha(x-x_i, y-y_i)$ of the model is much smaller than the change in the parameters. Therefore, when analyzing the parameter sensitivity, it can be regarded as a constant.

Assume that the relative error of the parameters caused by measurement is 1%. The wear distribution is taken as the result calculated by this model under the condition that the parameters have not changed, and the true wear amount d_{true} is taken as the simulation result in the model. After calculation, the model error is approximately linearly and positively correlated with the parameters, and the error change does not exceed $\pm 2.4\%$. It can be considered that this model is insensitive to parameter changes and relatively stable.

Model 2: Accelerated Wear Frequency Model

Sensitivity to the flow distribution pattern: In Model 2, we assume that the monthly pedestrian flow satisfies a certain probability distribution. Different probability distributions may affect the prediction error of the model. For the sensitivity analysis of the model to the distribution pattern, we studied the three most common distributions: uniform distribution, Gaussian distribution, and Poisson distribution.

After calculation, in Problem 4, the error generated by the uniform distribution estimate is 0.06, the error generated by the Gaussian distribution estimate is 0.0577, and the error generated by the Poisson distribution estimate is 0.0621. It can be seen that the error change is highly random and has no significant connection with the model used. It can be considered that the model is insensitive to the flow distribution pattern.

8 Strength and Weakness

8.1 Strength

- The use of genetic algorithms enables a more accurate description of step forces.
- A large amount of simulation data support makes the results more reliable and close to theoretical predictions.
- The gradient method and the regional method complement each other, making the prediction of renovation more reliable.

8.2 Weakness

- Some assumptions are simplistic and not sufficient to fully address real-world situations

References

Cho, Y. J., Lee, C., Lee, J. H., Kyung, M. G., Lee, K. H., & Lee, D. Y. (2021). The difference of in-shoe plantar pressure between level walking and stair walking in healthy males. *Journal of Biomechanics*, *122*, 110446.

Fujiyama, T., & Tyler, N. (2010). Predicting the walking speed of pedestrians on stairs. *Transportation Planning and Technology*, *33*(2), 177–202.

He, L., Fan, W., Dai, Z., Zhang, B., Li, Z., Guan, X., Xu, C., & Liu, T. (2023). A human-machine-environment interactive measurement system for non-anthropomorphic exoskeletons. *IEEE Transactions on Instrumentation and Measurement*.

Karaca, Z., Yılmaz, N. G., & Goktan, R. (2012). Abrasion wear characterization of some selected stone flooring materials with respect to contact load. *Construction and Building Materials*, *36*, 520–526.

Mohammed, A. A. A., Fener, M., Comakli, R., İnce, İ., Balci, M. C., & Kayabalı, K. (2021). Investigation of the relationships between basic physical and mechanical properties and abrasion wear resistance of several natural building stones used in turkey. *Journal of Building Engineering*, *42*, 103084.

Vesela, L. (2019). Staircase-dimensions of stair steps and their deviations of geometrical accuracy. *IOP Conference Series: Materials Science and Engineering*, *471*(2), 022012.

Letter

To: Archaeological team
From: Team #2501420 of 2025 MCM
Date: January 27th, 2025
Subject: A model for inferring building information based on the wear of stairs

We are currently conducting a highly significant research project that mainly focuses on leveraging the wear and tear of building staircases to deeply infer various key pieces of information about the buildings. As we advance this research, to ensure the accuracy and reliability of the model we are constructing and thus draw more scientific and rigorous research conclusions, we urgently need you to provide the following crucial data with your professional knowledge and rich experience.

Firstly, the model we've developed requires detailed data on staircase wear. This is one of the core data points for our research, and its accuracy is of utmost importance. We suggest that you use 3D laser scanning technology to conduct a comprehensive scan of the staircase. By doing so, you can obtain a high - precision 3D model of the staircase surface, enabling precise measurement of the wear depths at different parts of each step.

Estimating the usage frequency of the staircase is also a key part of our research. This requires you to take multiple factors into account. Regarding the function of the building, for example, if it is a religious building, the usage frequency of the staircase may increase significantly during specific religious festivals or events.

When it comes to historical information, if there are exact historical documents recording details such as the construction and renovation dates of the building, please provide the sources and specific content of these documents in detail. In the absence of clear written records, we hope you can determine the approximate historical period of the building based on your archaeological experience. By studying the technological characteristics employed in the building and synthesizing this information, you can infer the approximate age range of the building and explain the basis and reliability of your inference.

Records of building maintenance and renovation are of great significance for our accurate analysis of staircase wear. Determine whether the staircase materials were replaced during maintenance. If so, record the types, sources of the new materials, and the specific parts where the replacement occurred. Also, find out if the renovation involved changes to the staircase structure, and note the approximate time when these renovations took place.

The above - mentioned data play a decisive role in constructing a more accurate model and delving into the historical information behind buildings and staircases. If you can offer any speculations, please also elaborate on the above - mentioned indicators. We will combine theoretical calculations to evaluate the most reliable speculations. You are welcome to communicate with us at any time.

Thank you!

Sincerely yours,

Your friends