

Changing Tides: Momentum Swings in Tennis

Summary

Situations change rapidly in tennis, and the incredible swings are often attributed to **momentum**. We analyze momentum through Wimbledon 2023 men's matches data.

For Task 1, we develop a **Leverage Quantitative Momentum Capture Model**. Firstly, we consider the effect of short-term motivation and confidence boosting on serve percentage, and utilize the **Monte Carlo** algorithm to compute the winning probability(match/set/game) at each point. Then we define **Leverage** as how much winning or losing a ball changes the winning probability and the momentum as **exponential moving average** of leverage. Finally, the Wimbledon final match flow is visualized and the results are shown.

For Task 2, we use a statistical test to verify that "momentum" does play role in the match. Through the **run test**, we found that at a 90% confidence level, the proportion of non-random game scores in the given dataset accounts for 45.11% of the total, which indicates the swings in the game are not random. Furthermore, through the contingency table independence test, we determine, at a 90% confidence level, that swings in play and runs of success by one player are correlated to the momentum.

For Task 3, we establish the **Grid-Search Random Forest Swing Predictor Model** to forecast short-term swings in momentum. We select 43 indicators as inputs for the model to predict **momentum transition scenarios** (no transition of advantage, the advantage tilting towards one-self, the advantage transitioning to the opponent) in the next s points. The model achieved a 94% successful classification rate on the training set, demonstrating good classification performance. Utilizing GSRF, we can also identify that the factor most related to match fluctuations is the **opponent's total score**.

It is necessary to consider the top 15 factors in the GSRF model in terms of feature importance, so the four factors of dimensionality reduction into **Expenditure, Expenditure, Running Distance, and Game Situation Factor** are utilized by **factor analysis**. We suggest players focus on improving physical fitness, enhancing mental resilience, and paying special attention to break and net point scoring techniques.

For Task 4, by modifying the Monte Carlo simulation winning rule to **three sets to two wins**, we apply the model in the 2023 **Wimbledon women's singles final**, which verifies that the model is generalizable. It is worth noting that the most related factor of swing in women's tennis is different from that in men's singles, it is precise **the total running distance**, which may be related to the physical strength difference between men and women.

We conduct **sensitivity analysis** on the parameters of the short-term motivation effect in the Leverage Quantitative Momentum Capture Model and the parameters in the GSRF Swing Predictor Model, which are low in sensitivity and have high robustness.

Finally, a one-page memo with suggestions to tennis coaches is also produced to help coaches instruct players to control the pace of play.

Keywords: Monte Carlo Simulation, Runs Test, GSRF, Factor Analysis

Contents

1	Introduction	3
1.1	Background	3
1.2	Restatement and Analysis of the Problem	3
1.3	Overview of Our Work	4
2	Assumptions and Justification	5
3	List of Notation	5
4	Data Pre-processing	5
5	Task1: Captures the Flow of Play	6
6	Model Overview	6
7	Sub-model I : Adding Water Continuously	7
7.1	Model Establishment	8
7.1.1	Control Equations and Boundary Conditions	8
7.1.2	Definition of the Mean Temperature	9
7.1.3	Determination of Heat Transfer Capacity	9
8	Sub-model II: Adding Water Discontinuously	9
8.1	Heating Model	10
8.1.1	Control Equations and Boundary Conditions	10
8.1.2	Determination of Inflow Time and Amount	10
8.2	Standby Model	10
8.3	Results	10
8.3.1	Determination of Parameters	10
8.3.2	Calculating Results	10
9	Correction and Contrast of Sub-Models	10
9.1	Correction with Evaporation Heat Transfer	11
9.2	Contrast of Two Sub-Models	11
10	Model Analysis and Sensitivity Analysis	11
10.1	The Influence of Different Bathtubs	11
10.1.1	Different Volumes of Bathtubs	11
11	Strength and Weakness	12
11.1	Strength	12
11.2	Weakness	12
12	Further Discussion	12
	Reference	13
	Appendices	16
	Appendix A First appendix	16
	Appendix B Second appendix	16

1 Introduction

1.1 Background

Tennis matches are exciting and ever-changing. Players who seem to be dominating experience incredible swings from time to time. For example, a player may be in control of the match with a series of points, but at one rally, the opponent suddenly breaks and the situation is reversed. We often attribute this dramatic turn of events to "momentum".

"Momentum" heightens emotional intensity in the match, creating unpredictability and anticipation, making each moment charged and intense.

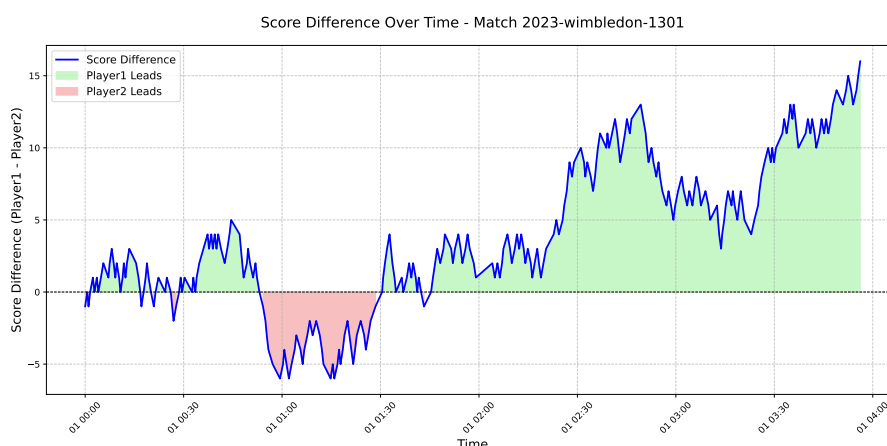


Figure 1: The fast-changing, heart-stopping swings of a tennis match.

1.2 Restatement and Analysis of the Problem

We need to complete the following tasks based on the data given in the question and additional data from ourselves:

- **Task 1: Build a model to capture the flow of play as points occur.** The model is required to be able to qualitatively and quantitatively describe performance of players at a given time and provide a visualization. In particular, the model should take the higher probability of the server winning into account.
- **Task 2: Assess whether swings in play and runs of success by one player during a match are random.**
- **Task 3: Look for indicators that identify the point in time when the game situation is reversed.**
 - Develop a model to predict swings in a match. Find the indicator with the highest correlation.
 - Based on momentum swings in past games, we need to give players advice on how to deal with other opponents.
- **Task 4: Evaluate the generalisability of the model we have developed to other competitions.**

- **Task 5: Write a memo summarising the results and giving targeted advice to coaches and players.**

After some analysis, we found that there is a close relationship between the tasks: In Task 1 and Task 2, we need to build a model describing the situation of the game (i.e. the relative performance of two players) and explore the relationship between "momentum" and changes in the game situation, respectively. The results of these two tasks are the basis of task 3. In task 3, we need to further build a model that can predict the situation fluctuations in the competition and find out the key factors that affect the competition. We also need to start to consider the applicability of our model. Can it provide competition suggestions for players to achieve better results? The fourth task is to consider the extensibility of our model. Are we still valuable in other competitions? Finally, in the memo, we need to focus on our model and its value.

In summary, we should effectively build a model that can predict the fluctuations of the competition situation, provide practical advice to players and coaches, and be extensible across different competitions.

1.3 Overview of Our Work

On the basis of the above analysis, we have carried out our work. Our working framework is shown in Figure 2.

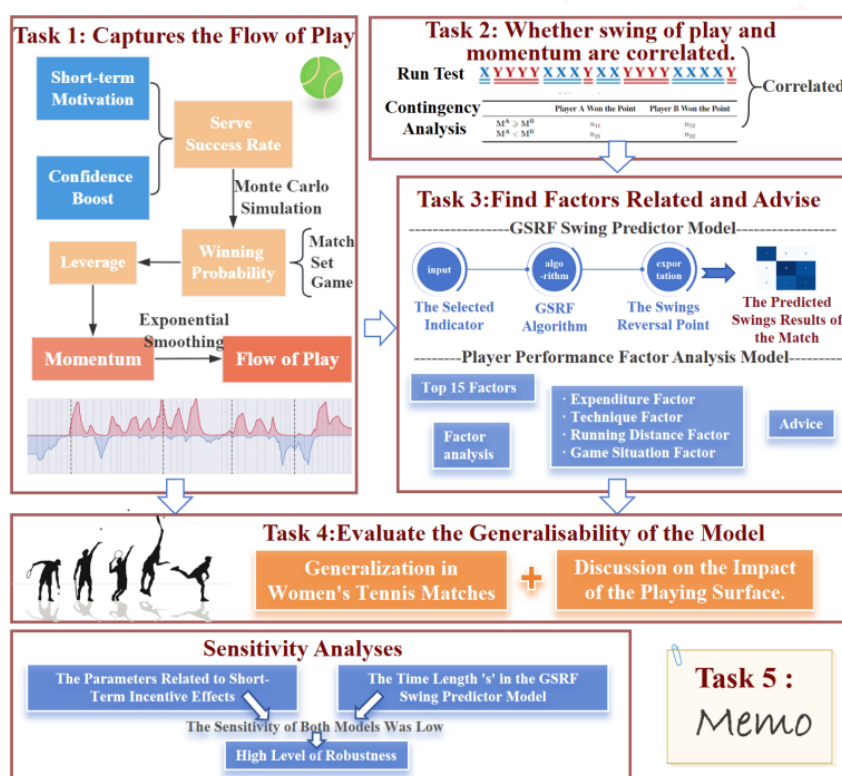


Figure 2: Our Work Overview Schematic Diagram

Figure 2: How To Draw?

The detailed modeling process and model details are presented in the following article.

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.

- **When a player wins a point, he experiences a short-term "motivational" effect that increases the probability of winning the next point.** The motivation effect can lead to increased self-confidence on a psychological level as well as increased physical conditioning on a physiological level.
- **Lifting effects are cumulative.** A string of successes may have a positive psychological and emotional effect on a player, so that a player winning more points in a row will result in a greater probability of winning the next point on serve.

3 List of Notation

Symbol	Meaning
p, q	Probability of Winning
m_i	Increase in the Rate of Serve points
T_{AB}	The Cumulative Total Score
L_i	Leverage of the i th Point
G	The leverage Gained
M	Momentum
λ	Tuning Parameter

Noted: Symbols not specified in the table should be referred to their first occurrence.

4 Data Pre-processing

• Outlier Handling

No clear outliers were observed in the data. However, in a match between Frances Tiafoe and Grigor Dimitrov during the third round, a notable occurrence occurred with a 24-hour time gap between two points at a set score of 2:0, as shown in Table 1.

Table 1: Suspected Outliers

match_id	playerA	playerB	elapsed_time
2023-wimbledon-1303	Frances Tiafoe	Grigor Dimitrov	0:53:55
2023-wimbledon-1303	Frances Tiafoe	Grigor Dimitrov	0:54:22
2023-wimbledon-1303	Frances Tiafoe	Grigor Dimitrov	24:56:34
2023-wimbledon-1303	Frances Tiafoe	Grigor Dimitrov	24:57:00

Upon consulting relevant information, we found the following explanation:

". Saturday's game was suspended due to inclement weather, so it was postponed to a day later to continue on Sunday.. . . . "

Therefore, the data is correct.

- **Missing Value Handling**

Missing Value Handling In the dataset given in the question, the quantities $serve_{width}$, $serve_{depth}$, and $return_{depth}$ have missing values, for which we use the **nearest-neighbour** filling method to fill them.

- **Tag Encoding of String Type Data**

There are many categorical variables in the data given in the title, presented as strings. In order to make it easier for the model to receive this type of data, we preprocess this type of data by means of label coding, using virtual coding for the last three columns of $serve_{width}$, $serve_{depth}$, and $return_{depth}$ in the data of the tennis match.

5 Task1: Captures the Flow of Play

Symbols	Description	Unit
h	Convection heat transfer coefficient	W/(m ² · K)
k	Thermal conductivity	W/(m · K)
c_p	Specific heat	J/(kg · K)
ρ	Density	kg/m ³
δ	Thickness	m
t	Temperature	°C, K
τ	Time	s, min, h
q_m	Mass flow	kg/s
Φ	Heat transfer power	W
T	A period of time	s, min, h
V	Volume	m ³ , L
M, m	Mass	kg
A	Aera	m ²
a, b, c	The size of a bathtub	m ³

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

6 Model Overview

In our basic model, we aim at three goals: keeping the temperature as even as possible, making it close to the initial temperature and decreasing the water consumption.

We start with the simple sub-model where hot water is added constantly. At first we introduce convection heat transfer control equations in rectangular coordinate system. Then we define the mean temperature of bath water.

Afterwards, we introduce Newton cooling formula to determine heat transfer capacity. After deriving the value of parameters, we get calculating results via formula deduction and simulating results via CFD.

Secondly, we present the complicated sub-model in which hot water is added discontinuously. We define an iteration consisting of two process: heating and standby. As for heating process, we derive control equations and boundary conditions. As for standby process, considering energy conservation law, we deduce the relationship of total heat dissipating capacity and time.

Then we determine the time and amount of added hot water. After deriving the value of parameters, we get calculating results via formula deduction and simulating results via CFD.

At last, we define two criteria to evaluate those two ways of adding hot water. Then we propose optimal strategy for the user in a bathtub. The whole modeling process can be shown as follows.

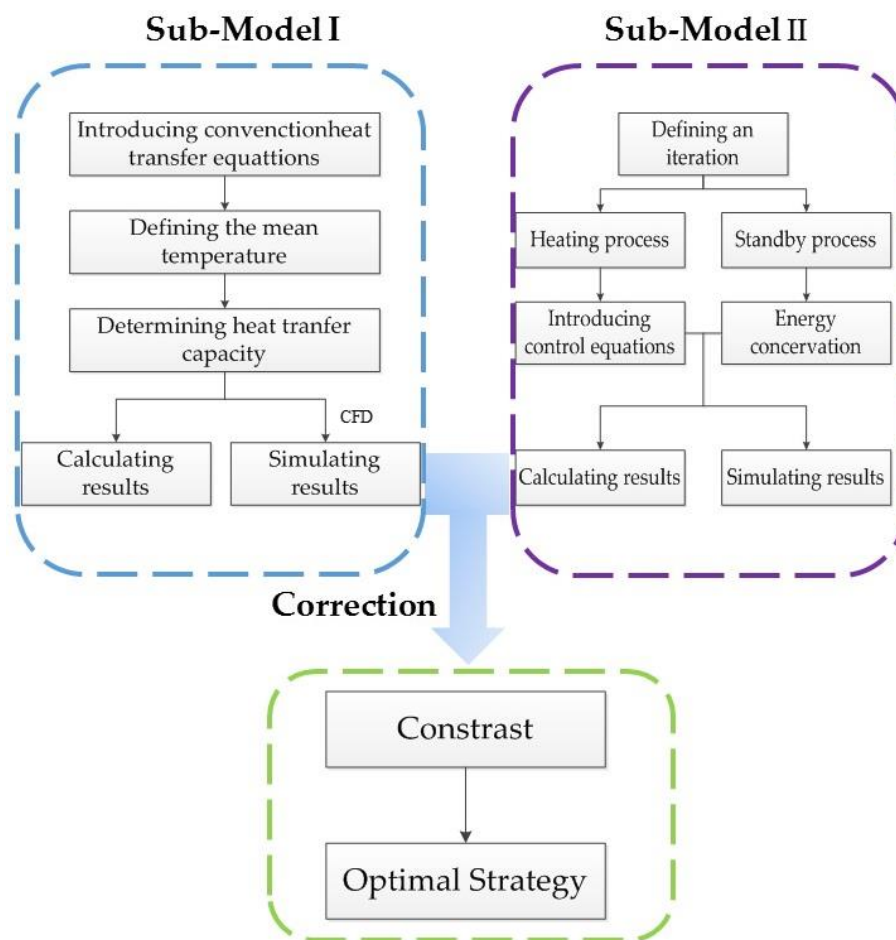


Figure 3: Modeling process

7 Sub-model I : Adding Water Continuously

We first establish the sub-model based on the condition that a person add water continuously to reheat the bathing water. Then we use Computational Fluid Dynamics (CFD) to simulate the change of water temperature in the bathtub. At last, we evaluate the model with the criteria which have been defined before.

7.1 Model Establishment

Since we try to keep the temperature of the hot water in bathtub to be even, we have to derive the amount of inflow water and the energy dissipated by the hot water into the air.

We derive the basic convection heat transfer control equations based on the former scientists' achievement. Then, we define the mean temperature of bath water. Afterwards, we determine two types of heat transfer: the boundary heat transfer and the evaporation heat transfer. Combining thermodynamic formulas, we derive calculating results. Via Fluent software, we get simulation results.

7.1.1 Control Equations and Boundary Conditions

According to thermodynamics knowledge, we recall on basic convection heat transfer control equations in rectangular coordinate system. Those equations show the relationship of the temperature of the bathtub water in space.

We assume the hot water in the bathtub as a cube. Then we put it into a rectangular coordinate system. The length, width, and height of it is a , b and c .

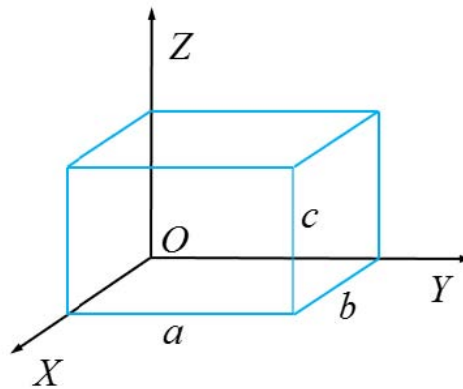


Figure 4: Modeling process

In the basis of this, we introduce the following equations [5]:

- **Continuity equation:**

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad (1)$$

where the first component is the change of fluid mass along the X -ray. The second component is the change of fluid mass along the Y -ray. And the third component is the change of fluid mass along the Z -ray. The sum of the change in mass along those three directions is zero.

- **Moment differential equation (N-S equations):**

$$\begin{cases} \rho \left(u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} \right) = -\frac{\partial p}{\partial x} + \eta \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) \\ \rho \left(u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} \right) = -\frac{\partial p}{\partial y} + \eta \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2} \right) \\ \rho \left(u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} \right) = -g - \frac{\partial p}{\partial z} + \eta \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} \right) \end{cases} \quad (2)$$

• **Energy differential equation:**

$$\rho c_p \left(u \frac{\partial t}{\partial x} + v \frac{\partial t}{\partial y} + w \frac{\partial t}{\partial z} \right) = \lambda \left(\frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} + \frac{\partial^2 t}{\partial z^2} \right) \quad (3)$$

where the left three components are convection terms while the right three components are conduction terms.

By Equation (3), we have

.....

On the right surface in Fig. 4, the water also transfers heat firstly with bathtub inner surfaces and then the heat comes into air. The boundary condition here is

7.1.2 Definition of the Mean Temperature

.....

7.1.3 Determination of Heat Transfer Capacity

.....

8 Sub-model II: Adding Water Discontinuously

In order to establish the unsteady sub-model, we recall on the working principle of air conditioners. The heating performance of air conditions consist of two processes: heating and standby. After the user set a temperature, the air conditioner will begin to heat until the expected temperature is reached. Then it will go standby. When the temperature get below the expected temperature, the air conditioner begin to work again. As it works in this circle, the temperature remains the expected one.

Inspired by this, we divide the bathtub working into two processes: adding hot water until the expected temperature is reached, then keeping this condition for a while unless the temperature is lower than a specific value. Iterating this circle ceaselessly will ensure the temperature kept relatively stable.

8.1 Heating Model

8.1.1 Control Equations and Boundary Conditions

8.1.2 Determination of Inflow Time and Amount

8.2 Standby Model

8.3 Results

We first give the value of parameters based on others' studies. Then we get the calculation results and simulating results via those data.

8.3.1 Determination of Parameters

After establishing the model, we have to determine the value of some important parameters.

As scholar Beum Kim points out, the optimal temperature for bath is between 41 and 45°C [1]. Meanwhile, according to Shimodozono's study, 41°C warm water bath is the perfect choice for individual health [2]. So it is reasonable for us to focus on 41°C ~ 45°C. Because adding hot water continuously is a steady process, so the mean temperature of bath water is supposed to be constant. We value the temperature of inflow and outflow water with the maximum and minimum temperature respectively.

The values of all parameters needed are shown as follows:

.....

8.3.2 Calculating Results

Putting the above value of parameters into the equations we derived before, we can get the some data as follows:

Table 2: The calculating results

Variables	Values	Unit
A_1	1.05	m^2
A_2	2.24	m^2
Φ_1	189.00	W
Φ_2	43.47	W
Φ	232.47	W
q_m	0.014	g/s

From Table 2,

.....

9 Correction and Contrast of Sub-Models

After establishing two basic sub-models, we have to correct them in consideration of evaporation heat transfer. Then we define two evaluation criteria to compare the two sub-models in order to determine the optimal bath strategy.

11 Strength and Weakness

11.1 Strength

- We analyze the problem based on thermodynamic formulas and laws, so that the model we established is of great validity.
- Our model is fairly robust due to our careful corrections in consideration of real-life situations and detailed sensitivity analysis.
- Via Fluent software, we simulate the time field of different areas throughout the bathtub. The outcome is vivid for us to understand the changing process.
- We come up with various criteria to compare different situations, like water consumption and the time of adding hot water. Hence an overall comparison can be made according to these criteria.
- Besides common factors, we still consider other factors, such as evaporation and radiation heat transfer. The evaporation turns out to be the main reason of heat loss, which corresponds with other scientist's experimental outcome.

11.2 Weakness

- Having knowing the range of some parameters from others' essays, we choose a value from them to apply in our model. Those values may not be reasonable in reality.
- Although we investigate a lot in the influence of personal motions, they are so complicated that need to be studied further.
- Limited to time, we do not conduct sensitivity analysis for the influence of personal surface area.

12 Further Discussion

In this part, we will focus on different distribution of inflow faucets. Then we discuss about the real-life application of our model.

- Different Distribution of Inflow Faucets

In our before discussion, we assume there being just one entrance of inflow.

From the simulating outcome, we find the temperature of bath water is hardly even. So we come up with the idea of adding more entrances.

The simulation turns out to be as follows

From the above figure, the more the entrances are, the evener the temperature will be. Recalling on the before simulation outcome, when there is only one entrance for inflow, the temperature of corners is quietly lower than the middle area.

In conclusion, if we design more entrances, it will be easier to realize the goal to keep temperature even throughout the bathtub.

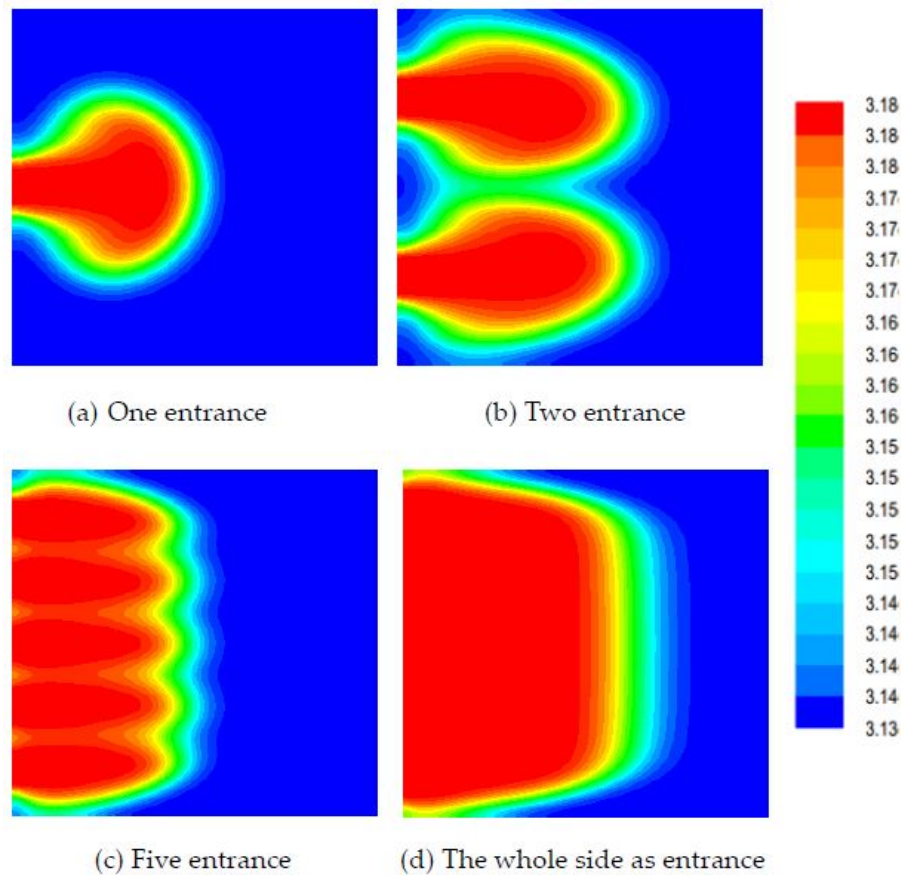


Figure 5: The simulation results of different ways of arranging entrances

• Model Application

Our before discussion is based on ideal assumptions. In reality, we have to make some corrections and improvement.

- 1) Adding hot water continually with the mass flow of 0.16 kg/s. This way can ensure even mean temperature throughout the bathtub and waste less water.
- 2) The manufacturers can design an intelligent control system to monitor the temperature so that users can get more enjoyable bath experience.
- 3) We recommend users to add bubble additives to slow down the water being cooler and help cleanse. The additives with lower thermal conductivity are optimal.
- 4) The study method of our establishing model can be applied in other area relative to convection heat transfer, such as air conditioners.

References

- [1] Gi-Beum Kim. Change of the Warm Water Temperature for the Development of Smart Healthcare Bathing System. Hwahak konghak. 2006, 44(3): 270-276.
- [2] https://en.wikipedia.org/wiki/Convective_heat_transfer#Newton.27s_law_of_cooling
- [3] https://en.wikipedia.org/wiki/Navier%E2%80%93Stokes_equations

-
- [4] https://en.wikipedia.org/wiki/Computational_fluid_dynamics
 - [5] Holman J P. Heat Transfer (9th ed.), New York: McGraw-Hill, 2002.
 - [6] Liu Weiguo, Chen Zhaoping, ZhangYin. Matlab program design and application. Beijing: Higher education press, 2002. (In Chinese)

Enjoy Your Bath Time!

From simulation results of real-life situations, we find it takes a period of time for the inflow hot water to spread throughout the bathtub. During this process, the bath water continues transferring heat into air, bathtub and the person in bathtub. The difference between heat transfer capacity makes the temperature of various areas to be different. So that it is difficult to get an evenly maintained temperature throughout the bath water.

In order to enjoy a comfortable bath with even temperature of bath water and without wasting too much water, we propose the following suggestions.

- Adding hot water consistently
- Using smaller bathtub if possible
- Decreasing motions during bath
- Using bubble bath additives
- Arranging more faucets of inflow

Sincerely yours,

Your friends

Appendices

Appendix A First appendix

In addition, your report must include a letter to the Chief Financial Officer (CFO) of the Goodgrant Foundation, Mr. Alpha Chiang, that describes the optimal investment strategy, your modeling approach and major results, and a brief discussion of your proposed concept of a return-on-investment (ROI). This letter should be no more than two pages in length.

Here are simulation programmes we used in our model as follow.

Input matlab source:

```
function [t,seat,aisle]=OI6Sim(n,target,seated)
pab=rand(1,n);
for i=1:n
    if pab(i)<0.4
        aisleTime(i)=0;
    else
        aisleTime(i)=trirnd(3.2,7.1,38.7);
    end
end
end
```

Appendix B Second appendix

some more text **Input C++ source:**

```
//=====
// Name       : Sudoku.cpp
// Author      : wzlf11
// Version     : a.0
// Copyright   : Your copyright notice
// Description : Sudoku in C++.
//=====

#include <iostream>
#include <cstdlib>
#include <ctime>

using namespace std;

int table[9][9];

int main() {

    for(int i = 0; i < 9; i++){
        table[0][i] = i + 1;
    }

    srand((unsigned int)time(NULL));

    shuffle((int *)&table[0], 9);

    while(!put_line(1))
```



```
{
    shuffle((int *)&table[0], 9);
}

for(int x = 0; x < 9; x++){
    for(int y = 0; y < 9; y++){
        cout << table[x][y] << " ";
    }

    cout << endl;
}

return 0;
}
```

Report on Use of AI

1. OpenAI ChatGPT (Nov 5, 2023 version, ChatGPT-4,)

Query1: <insert the exact wording you input into the AI tool>

Output: <insert the complete output from the AI tool>

2. OpenAI Ernie (Nov 5, 2023 version, Ernie 4.0)

Query1: <insert the exact wording of any subsequent input into the AI tool>

Output: <insert the complete output from the second query>

3. Github CoPilot (Feb 3, 2024 version)

Query1: <insert the exact wording you input into the AI tool>

Output: <insert the complete output from the AI tool>

4. Google Bard (Feb 2, 2024 version)

Query1: <insert the exact wording of your query>

Output: <insert the complete output from the AI tool>