Information Dynamics in Baseball

Takeo Nakagawa

Satellite College Hakusan

npo.hakusan@kjc.biglobe.ne.jp

Abstract

This study is concerned with the information dynamics in baseball. This simulates and models the history of professional baseball game Carp vs. Giants in Japan Central League , and provides us useful strategy and/or tactic for coping with the future games, and method for predicting the game outcomes. The game, Carp vs. Giants is quite a unique one, for Giants keeps the advantage against Carp over 90% of the total game length, but is reversed by Carp just before the end. Thus, this game should be called as one-sided game by Giants, though Carp becomes the winner. It is found that a model curve for height, which initial slope angle coincides with that of data curve, provides the minimum current value of height depending on the time. Usefulness of the information dynamic model for predicting game outcome has been demonstrated: It is predicted in such a way that the author's granddaughter in 10 years old on August 13, 2018 will increase her height in 10 years from 127.5 cm to 167.6 cm.

Keywords: Game Theory, Baseball, Advantage, Certainty of Game Outcome, Prediction of Game Outcome, Game Point

1. Introduction

Since 2010, a new game theory of information dynamics has been researched extensively and reported (Iida et al. 2012a, Nakagawa and Minatoya 2014). This theory is a breakthrough in game research, for it incorporates notion of the time in the game analyses for the first time. The theory is primarily aimed to clarify how certainty of game outcome, varies with time before, during and after the game. In principle, we could not only reflect the game after it is over, but also predict the game history and outcome before the start, if necessary and sufficient information regarding to players (or teams), is provided. So far, the theory has been applied to Judo (Nakagawa and Minatoya 2014), baseball (Iida et al. 2012a), effect of medicine (Iida et al. 2012b), Soccer(Nakagawa and Nakagawa 2014) or Shogi(Nakagawa et al. 2014) successfully.

As far as the present author is aware of, no previous theory has succeeded to analyze a game in such a way that it varies with game length (or time). For example, von Neumann's game theory (Neuman and Morgenstern 1944) is concerned with only a few outcomes due to each the player's decision without considering their time dependency.

Main purpose of the present study is to simulate and model the history and outcome of professional baseball game in the Japan Central League, Carp vs. Giants held at Matsuda Stadium, Hiroshima, Japan on July 19 in 2019 by data analyses together with information dynamic model. The present work is aimed at providing readers useful strategy and/or tactic for coping with the future games, and for predicting the game outcomes.

2. Method of Analyses

For clarity, elemental procedures for obtaining the advantage α, certainty of game outcome ξ, and uncertainty of game outcome ς will be explained by using soccer between teams A and B, where only goal is assumed to be the evaluation function score, which is defined as decisive factor in game

The advantage α is defined as follows: When the total score(s) of the two teams at the end of game $S_T \neq 0$,

$\alpha = [S_A(\eta) - S_B(\eta)]/S_T$ for $0 \leq \eta \leq 1$, (1)

where $S_A(\eta)$ is the current score sum for team A(winner), $S_B(\eta)$ is the current score sum for team B(loser) and η is the normalized game length(or time) This means that when $\alpha>0$, team A (winner) gets the advantage against team B (loser) in the game, while when $\alpha<0$, team B (loser) gets the advantage against team A (winner). It is certain that when $\alpha=0$ the game is balanced. It is critical how assessor(s) chooses and assesses the evaluation function scores during the game properly in order to obtain useful results of analyses. When the total score(s) of the two teams at the end of game $S_T=0$, $\alpha = 0$ for $0 \leq \eta \leq 1$.

The certainty of game outcome ξ during the game is defined as follows: When the total score(s) of the two teams at the end of game $S_T \neq 0$,

$\xi = |S_A(\eta) - S_B(\eta)|/S_T$ for $0 \leq \eta < 1$

$=1$(normal game) or 0 (draw game) for $\eta=1$. (2)

At $\eta=1$, ξ is assigned to the value of 1, for at the end of game the information on the game outcome must be 100%. The reason why we take the absolute value of the advantage α to get the certainty of game outcome ξ for $0 \leq \eta<1$ is that ξ is independent of the sign of α. This may be reasonable if one consider meaning of the certainty of game outcome: As far as the absolute value of the advantage α increases (decreases), the certainty of game outcome ξ must increase (decrease). In case of draw game, ξ may be assigned to the value of 0 at $\eta=1$, which is end of the game. When the total score(s) of the two teams at the end of game $S_T= 0$, $\xi= 0$ for $0 \leq \eta\leq 1$.

The uncertainty of game outcome ς during the game is defined as follows

$\varsigma = 1 - \xi$. (3)

This equation denotes that the current uncertainty of game outcome ς can be obtained by subtracting the current certainty of game outcome ξ from value of 1.

The game length is the current length (or time) from the beginning of the game, and is normalized by the total game length (or total time) to obtain the normalized game length η.

3. Results

In this section, Carp vs. Giants has been analyzed, and are summarized in Table 1. In this analysis, the evaluation function score is the game score, and each the score is assigned as value of 1. The total scores S_T is value of 13, and the total time is 18 innings. In Table 1, records and results of data analyses in game, Carp vs. Giants are summarized. Giants leads the score until η=0.889(the first of 8th inning), but Carp reverses the score at η=0.944(the second of 8th inning), where Carp, Aizawa makes 2 run homerun into the right stand by hitting the outside straight ball thrown by Giants, Majison. At the first of 9th inning, Giants gets no score, so that the game ends and Carp wins the game by the score of 7 : 6

Table 1 Records and results of data analyses in game, Carp vs. Giants

Inning	η	S_A	S_B	α	ξ
0	0	0	0	0	0
1 first	0.111	0	1	-0.077	0.077
1 second	0.167	0	1	-0.077	0.077
2 first	0.222	0	1	-0.077	0.077
2 second	0.278	0	1	-0.077	0.077
3 first	0.333	0	2	-0.154	0.154
3 second	0.389	0	2	-0.154	0.154
4 first	0.444	0	5	-0.385	0.385
4 second	0.5	1	5	-0.308	0.308
5 first	0.556	1	5	-0.308	0.308
5 second	0.611	2	5	-0.231	0.231
6 first	0.667	2	5	-0.231	0.231
6 second	0.722	3	5	-0.154	0.154
7 first	0.778	3	5	-0.154	0.154
7 second	0.833	4	5	-0.077	0.077

8 first	0.889	4	6	-0.154	0.154
8 second	0.944	7	6	0.077	0.077
9 first	1	7	6	0.077	1

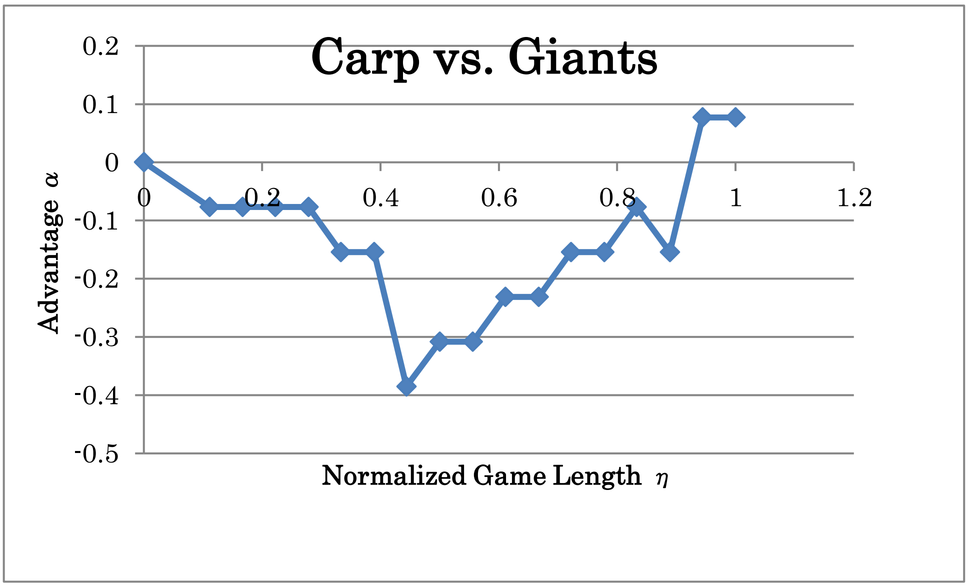

Figure 1 Advantage α against normalized game length η

Figure 1 shows the relation between the advantage α and the normalized game length η. This figure shows that starting from the value of 0, α is negative until η=0.889, and takes the minimum value of -0.385 at η=0.444. However, at η=0.944, α become positive value of 0.077, and it continues until the end. This means that Giants keeps the advantage against Carp over 90% of the total game length, though Giants loses the win in this game.

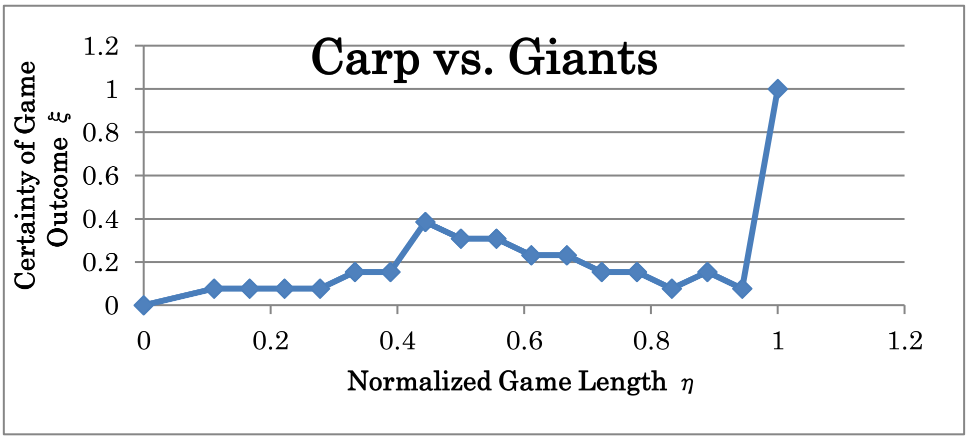

Figure 2 Certainty of game outcome ξ against normalized game length η

Figure 2 shows the relation between the certainty of game outcome ξ and the normalized game length η. This figure shows that after starting from 0, ξ increases to 0.385 at η=0.444 with increasing η, then decreases to 0.077 at η=0.944 and jumps to the value of 1 in the end..

4. Discussion

4.1 Game Point

The game point of Carp vs. Giants, which is the cross point between the certainty of game outcome ξ and the uncertainty of game outcome ς will be discussed in terms of information dynamic model,

$\xi=\eta^m$, (4)

where ξ is the certainty of game outcome, η is the normalized game length, and m is the positive real number. For the details about the information dynamic model, refer to Appendix.

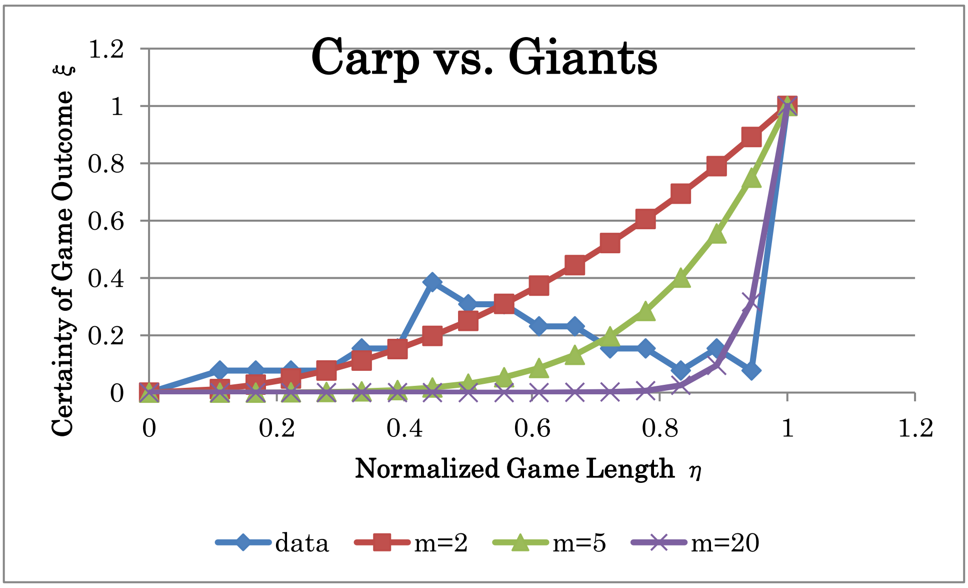

Figure 3 Certainty of game outcome ξ against normalized game length η

Figure 3 shows the relation between the certainty of game outcome ξ and the normalized game length η, where the game data, Carp vs. Giants, and the three curves of information dynamic model $\xi=\eta^m$, m=2, 5, and 20, respectively have been plotted concurrently. The data curve follows fairly well to the model curve $\xi=\eta^2$ until η=0.4, but it follows well to the model curve $\xi=\eta^{20}$ after η=0.85. It may be evident that the model curve change from $\xi=\eta^2$ to $\xi=\eta^{20}$ is caused by the Carp' score of 1 at 4th, 5th, 6th, and 7th innings, respectively. It is, therefore, interesting to discuss the two game points; the first one is the cross point between $\xi=\eta^2$ and $\varsigma=1-\eta^2$, and the second one is the cross point between $\xi=\eta^{20}$ and $\varsigma=1-\eta^{20}$.

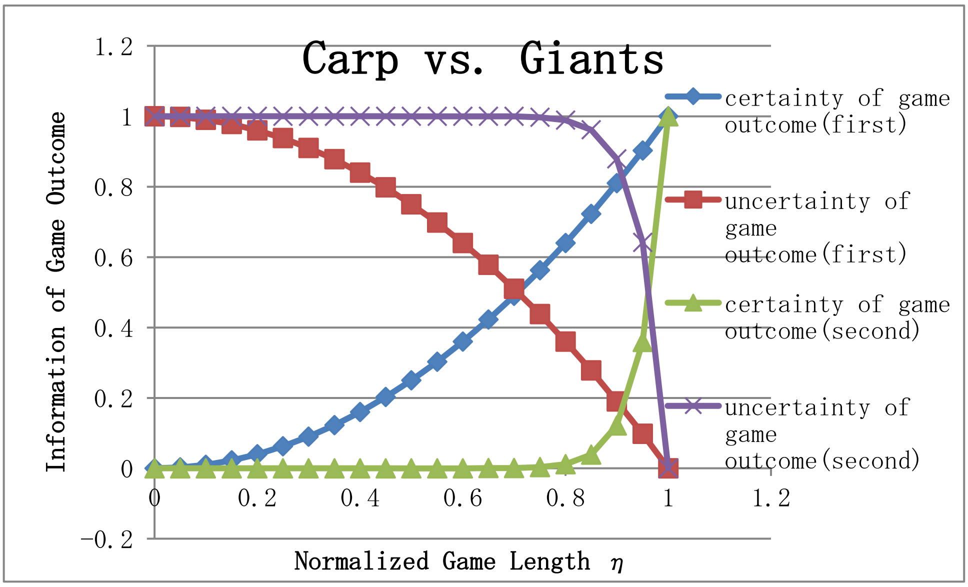

Figure 4 Information of game outcome against normalized game length η

Figure 4 shows the relation between the information of game outcome and the normalized game length η. The first cross point between the certainty of game outcome(first) and the uncertainty of game outcome(first) is the game point of ca. η=0.7. That is, if the game proceeds following the model $\xi=\eta^2$ in keeping Giants advantage against Carp throughout the game , Giants must get the safety lead at this stage. On one hand, the second cross point between the certainty of game outcome (second) and the uncertainty of game outcome(second) is the game point of ca. η=0.97. That is, if the game proceeds following the model $\xi=\eta^{20}$ in keeping Carp advantage against Giants throughout the game, Carp must get the safety lead at this stage. Indeed, this game is quite a unique one, for Giants keeps the advantage over 90% of the total game length, but is reversed by Carp at just before the end at η=0.944. Thus, this can be said to be 'one-sided game' by Giants, but Carp becomes the winner by reversing the score just before the end finally.

4.2 Height and Information Dynamic Model

Let us discuss the relation between the height and the information dynamic model based on the data of my colleague, Ryusui Masakakio from age 12 to 22. Table 2 shows the records and results of data analyses on height increment of Ryusui Masakakio. In this Table, Δh is the height increment from 150 cm depending on his age, η is the time(or game length) normalized by the total time of 10 years and $\hat{h}$ is the height increment normalized by the total height increment of 25.5 cm. His height at age 12 is 150 cm, and the initial height increment at this age is 8 cm/year.

Table 2 Records and results of data analyses on height increment of Ryusui Masakakio. Age h Δh η ĥ

12	150	0	0	0
13	158	8	0.1	0.313
14	164	14	0.2	0.549
15	169	19	0.3	0.745
16	170	20	0.4	0.784
17	171	21	0.5	0.824
18	172	22	0.6	0.863
19	173	23	0.7	0.902
20	174	24	0.8	0.941
21	175	25	0.9	0.98
22	175.5	25.5	1	1

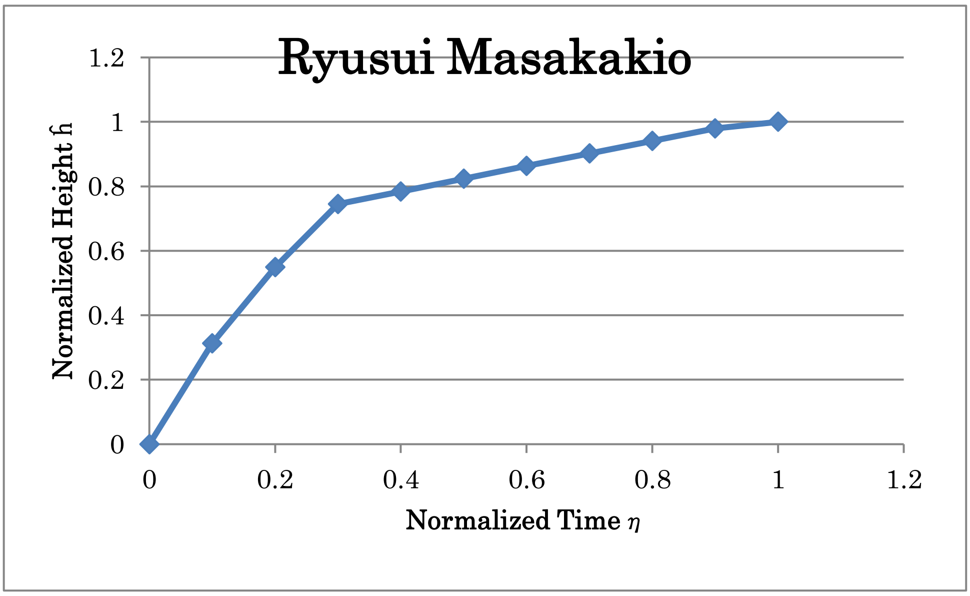

Figure 5 Normalized height ĥ against normalized time η

Figure 5 shows the relation between the normalized height ĥ and the normalized time η.

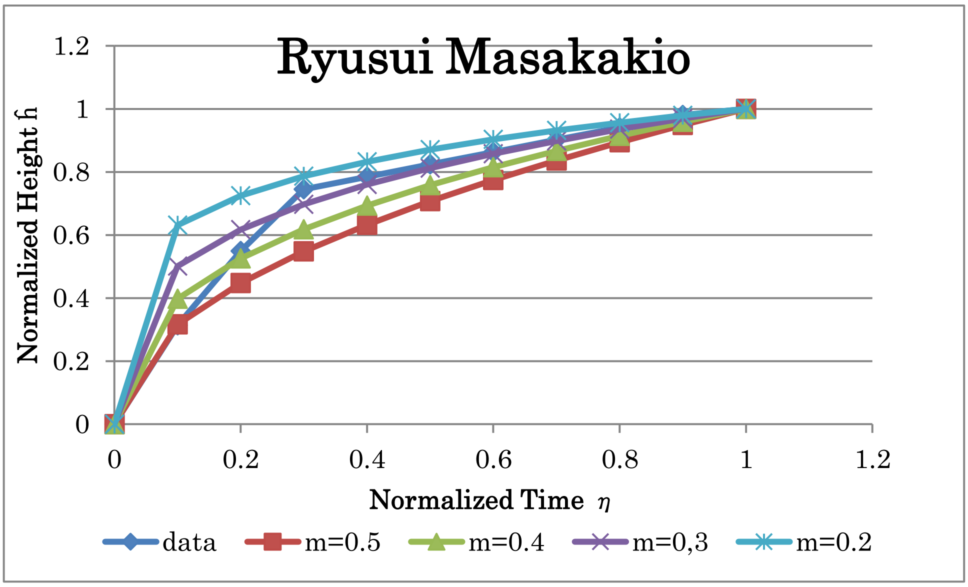

Figure 6 Normalized height ĥ against normalized time η

Figure 6 shows the relation between the normalized height ĥ and the normalized time η, where the data curve and four model curves $\xi = \eta^m$, where m=0.5, 0.4, 0.3, and 0.2, respectively, are plotted concurrently. It is worth noting that the initial slope angle of the data curve coincides with that of the model curve $\xi=\eta^{0.5}$, but the data curve increases with η more rapidly than this model curve. The data curve coincides with the model curve $\xi=\eta^{0.5}$ until η=0.1, but deviates from this model curve with increasing η. Then, the data curve approaches and merges to the model curve $\xi=\eta^{0.3}$ at η=0.5, and follows this curve until the end. It is realized that the current value of ĥ is an intermediate value between the model curve $\xi=\eta^{0.5}$ and 1. Thus, it is considered that the model curve $\xi=\eta^{0.5}$, which initial slope angle coincides with that of the data curve, provides the minimum current value of height. This knowledge is useful to predict the height depending on the time (or game length) in terms of information dynamic model.

4.3 Prediction of Height for a Girl

It may be interesting to point out that the information dynamic model possesses a potential to predict game outcome. Using initial conditions such as team (player)' ranking, record or quality of players, value of the parameter 'm' of the information dynamic model $\xi=\eta^m$ may be obtained before a game starts. It is clear that once this value is provided, the game will proceed along one of the model curves in Figure 9 from the start to the end. The information dynamic model is also applicable to predict future trends in social problems such as GDP (Gross Domestic Product), population, temperature, effect of medicine, or achievement quotient of pupils in school, where current information such as the slope angle of the relevant curves is critical. It is worth noting that this approach is applicable to any serious game, for it supports learning and/or training for educator, nurse and many others.

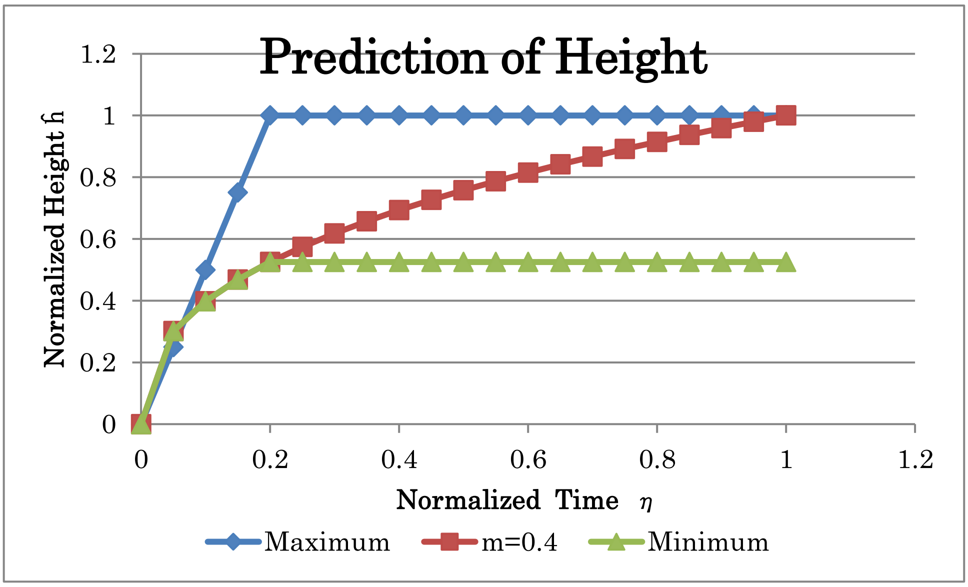

Figure 7 Normalized height ĥ against normalized time η

Let us illustrate how to predict the height of the author's granddaughter, Kinue Kanzaki in 10 years old. Her height is 127.5 cm on August 13, 2018, and increases 5.25 cm/year for the last two years, which provides the slope angle of the model curve at the start. The assumptions in this analysis are as follows,

(a) Her height only increases for 10 years from August 13, 2018, until August 12, 2028.

(b) Her life time is 60 years, so she will survive for 50 years from August 13, 2018 to August 12, 2068.

(c) The information dynamic model having the same slope angle as that of data curve, gives the minimum height at each time for 10 years from August 13, 2018 until August 12, 2028.

It is clear that all of the above assumptions are hypotheses, which will be revealed to be true or fault on August 12, 2068. The assumption (c) seems to be supported by results of analyses in section 4.2.

Figure 7 shows the relation between the normalized height ĥ and the normalized time η, where the height is normalized by the total height of 52.5 cm, which is total amount of the increase in 10 years if the initial increase rate of height is kept constant, while the time is normalized by the total time of 50 years. In case of the line 'Maximum', normalized height ĥ increases linearly from 0 to 1(or total height 52.5 cm) when the normalized time η changes from 0 to 0.02(or 10 years). Hence, the line 'maximum' gives the maximum height of 180(127.5+52.5) cm on August 12, 2028. However, it is certain that her height never increases at constant rate, but must increase the lesser rate with time, but the rate itself is unknown and depends on the individual, so that it is assumed here that an information dynamic model curve provides the minimum height: Plotted is the model curve $h=\eta^{0.4}$, in which the initial increase rate of 5(or 5.25 cm/year) coincides with that of the data curve. This might give the minimum height of 155.1(127.5+27.6) cm on August 12, 2028. It may be, therefore, expected that her height on August 12, 2028 might be in between 155.1 cm and 180 cm, so that her height is predicted to be about 167. 6 cm by taking the average. At the moment, the result itself is merely a conjected value, but this clearly demonstrates usefulness of the information dynamic model for the prediction of game outcome. Note that the 'Minimum' curve follows the model until the normalized time η=0.2(or 10 years), but keeps a constant value of the **normalized height $\hat{h}$=0.525 (or 27.6 cm), for it is assumed that once the** normalized time η exceeds this value of 0.2, increase of her height stops.

5. Conclusions

The new knowledge and insights obtained through the present study have been summarized as follows.

1. This study simulates and models the history of professional baseball game Carp vs. Giants , and provides us useful strategy and/or tactic coping with the future games, and method for predicting the game outcome.

2. The game, Carp vs. Giants is quite a unique one, for Giants keeps the advantage against Carp over 90% of the total game length, but is reversed by Carp just before the end. Thus, this game should be called as one-sided game by Giants, though Carp becomes the winner.

3. It is found that an information dynamic model curve for height, which initial slope angle coincides with that of data curve, provides the minimum current value of height depending on the time.

4. Usefulness of the information dynamic model for predicting game outcome has been demonstrated: It is predicted in such a way that the author's granddaughter in 10 years old on August 13, 2018 will increase her height in 10 years from 127.5 cm to 167.6 cm.

References

1. Iida, H., Nakagawa, T., Spoerter, K. (2012) Game information dynamic models based on fluid mechanics. Entertainment and Computing, 3, 89-99.

2. Iida, H., Nakagawa, T., Nossal, N. (2012) Certainty of patient survival with reference to game information dynamic model. Open Journal of Preventive Medicine, 2, 490-498.

3. Nakagawa, T.R.M., Minatoya, H. (2014) Information dynamics in judo. Research Journal of Budo, 47, 29-45.

4. Nakagawa, T., Nakagawa, A. (2014) A safety lead curve and entertainment in games. International Review of Social Sciences and Humanities, 7, 91-103.

5. Nakagawa, T.R.M., Iida, H. (2014) Three game patterns. International Journal of Research Studies in Computer Sciences and Engineering, 1, 1-12.

6. Nakagawa, T., Iida, H., Morichika, H., Ohura, S., Yonenaga, K. (2014) Human and computer interaction in shogi. Research Journal of Computation and Mathematics, 2, 7-13.

7. Neumann, J. von, Morgenstern, O. (1944) Theory of Games and Economics Behavior. Princeton University Press.

8. Blasius, H. (1907) Grenzschiten Flüssigkeiten mit kleiner Reibung. Inaugural-Dissertation zur Erlangung der Doktorwürde der Hohen Philosophischen Fakultät Georgia Augusta zur Göttingen, Druck von B.G. Teubner in Leipzig.

9. Prandtl L. (1904) Über Flüssigkeitsbewegung bei sehr kleiner Reibung. Verhandlungen des 3rd Internationalen Mathematiker-Kongresses, Heidelberg, Leipzig: Teubner.

10. Schlichting H. (1968) Boundary-Layer Theory (translated by J. Kestin), McGraw-Hill, New York, 6th Edition.

11. Solso, R. (1994) Cognition and the Visual Arts. MIT Press, Cambridge.

Appendix: Information Dynamic Model

Currently, information dynamic model only make it possible to treat and identify game progress patterns depending the game length (or time). In this model, certainty of game outcome is expressed as a simple analytical function depending on the game length. In this Appendix, the information dynamic model has been introduced in brief.

Modelling Procedure

The modeling procedure of information dynamics based on fluid mechanics is summarized as follows:

(a) Assume a flow problem as the information dynamic model and solve it.

(b) Get the solutions, depending on the position (or time).

(c) Examine whether any solution of the problem can correspond to game information.

(d) If so, visualize the assumed flow with some means. If not, return the first step.

(e) Determine the correspondence between the flow solution and game information.

(f) Obtain the analytical expression of the information dynamic model.

The information dynamic model will be constructed by following the above procedure step by step.

(a) Let us assume flow past a flat plate at incident angle of zero as the information dynamic model (Figure 8).

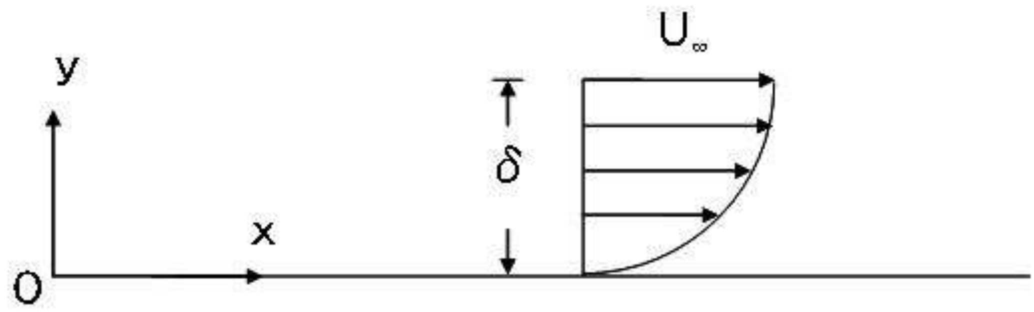

Figure 8 Definition sketch of flow past a flat plate at incident angle of zero.

An example of the application of the boundary-layer equations, which is the simplified version of Navier-Stokes equations(Schlichting 1968), is afforded by the flow along a very thin flat plate at incident angle of zero. Historically this is the first example illustrating the application of Prandtl's boundary layer theory (Prandtl 1904); it has been discussed by Blasius(1907) in his doctoral thesis at Göttingen. Let the leading edge of the plate be at x=0, the plate being parallel to the x-axis and infinitely long downstream, as shown in Figure 7. We shall consider steady flow with a free-stream velocity U, which is parallel to the x-axis. The boundary-layer equations (Prandtl 1904, Schlichting 1968) are expressed by

$u \cdot \partial u/\partial x + v \cdot \partial u/\partial y = -1/\rho \cdot dp/dx + \nu \, \partial^2 u/\partial y^2$, (5)

$\partial u/\partial x + \partial v/\partial y = 0$ (6)

$y = 0 : u = v = 0; \; y = 1 : u = U, v=0,$ (7)

where u and v are velocity components in the x- and y- directions, respectively, ρ the density, p the pressure and ν the kinematic viscosity of the fluid. In the free stream,

$U \cdot dU/dx = -1/\rho \cdot dp/dx.$ (8)

The free-stream velocity U is constant in this case, so that dp/dx=0, and dp/dy =0. Since the system under consideration has no preferred length it is reasonable to suppose that the velocity profiles at varying distances from the leading edge are similar to each other, which means that the velocity curves u(y) for varying distances x can be made identical by selecting suitable scale factors for u and y. The scale factors for u and y appear quite naturally as the free-stream velocity, U and the boundary-layer thickness, δ(x). Hence, the velocity profiles in the boundary-layer can be written as

$u/U = f(y/\delta).$ (9)

Blasius(1907) has obtained the solution in the form of a series expansion around y/δ= 0 and an asymptotic expansion for y/δ being very large, and then the two form being matched at a suitable value of y/δ.

(b)The similarity of velocity profile is here accounted by assuming that function f depends on y/δ only, and contains no additional free parameter. The function f must vanish at the wall (y = 0) and tend to the value of 1 at the outer edge of the boundary-layer (y = δ).

When using the approximate method, it is expedient to place the point at a finite distance from the wall, or in other words, to assume a finite boundary-layer thickness δ(x) , in spite of the fact that all exact solutions of boundary-layer equations tend asymptotically to the free-stream associated with the particular problem.

The "Approximate method" here means that all the procedures are to find approximate solutions to the exact solutions, respectively. When writing down an approximate solution of the present flow, it is necessary to satisfy certain boundary condition for u(y). At least the no-slip condition u = 0 at y = 0 and the condition of the continuity when passing from the boundary-layer profile to the free-stream velocity, so that u = U at y =δ, must be satisfied.

It is evident that the following velocity profile satisfies all of the boundary conditions for the assumed flow past a flat plate at incident angle of zero,

$u/U = (y/\delta)^m$, (10)

where m is positive real number. Eq.(10) is heuristically derived, and represents a group of the approximate solutions for the assumed flow, taking each the different value of m. In case of m=1, (10) reduces to an exact solution for the boundary-layer equations, and the rest solutions are considered as the approximate solutions to the other exact solutions, respectively.

(c) Let us examine whether these solutions are game information or not. Such an examination immediately provides us that the non-dimensional velocity varies from 0 to 1 with increasing the non-dimensional vertical distance y/δ in many ways as the non-dimensional information, so that these solutions can be game information. However, validity of this conjecture shall be confirmed by the relevant data.

(d) Imagine that the assumed flow is visualized with neutral buoyant particles. Motion of the visualized particles is detected by the eye almost instantaneously through the lights and is mapped on our retina first (Solso 1994), so that during these processes, motion of "fluid particles" is transformed into that of the "information particles" by the light carrying the images of fluid particles. This is why motion of the fluid particles is intact in the physical space, but only the reflected lights, or electromagnetic waves consisting of photons can reach at the retina. The photons are then converted to electrochemical particles and are passed along the visual cortex for further processing in parts of the cerebral cortex (Solso 1994). The photons and/or electrochemical particles are considered to be information particles. It is, therefore, natural to expect that the flow in the physical space is faithfully transformed to that in the information space, or brain including eye. During this transformation, the flow solution in the physical space changes into the information solution in the information space.

(e) Proposed are correspondences between the flow and game information, which are listed in Table 3.

Table 3 Correspondences between the flow and game information

Flow	Game
u: flow velocity	I:current information
U: free stream velocity	I_0: total information
y: vertical co-ordinate	t: current game length
δ: boundary layer thickness	t_0: total game length

(f) Considering the correspondences in Table 3, (10) can be rewritten as

$I/I_0 = (t/t_0)^m$ (11)

Introducing the following normalized variables in (11),

$\xi = I/I_0$ and $\eta = t/t_0$, (12)

we finally obtain the analytical expression of the information dynamic model as

$\xi = \eta^m$ (13)

where ξ is the certainty of game outcome, η is the normalized game length, and m is a positive real number.

Figure 9 shows the relation between certainty of game outcome ξ and normalized game length η (or time), where a total of 10 model curves have been plotted concurrently. This figure clearly suggests versatility of this model (13), for each of the curve represents one game being different from all of the other games. Thus, this model can represent any game in principle, for the parameter 'm' can take any positive real number The smaller the strength difference between both players (or teams) is, the greater the value of m, and *vice versa*. This means that each the game takes a particular value of m, and thus experiences its own unique history with increasing the time.

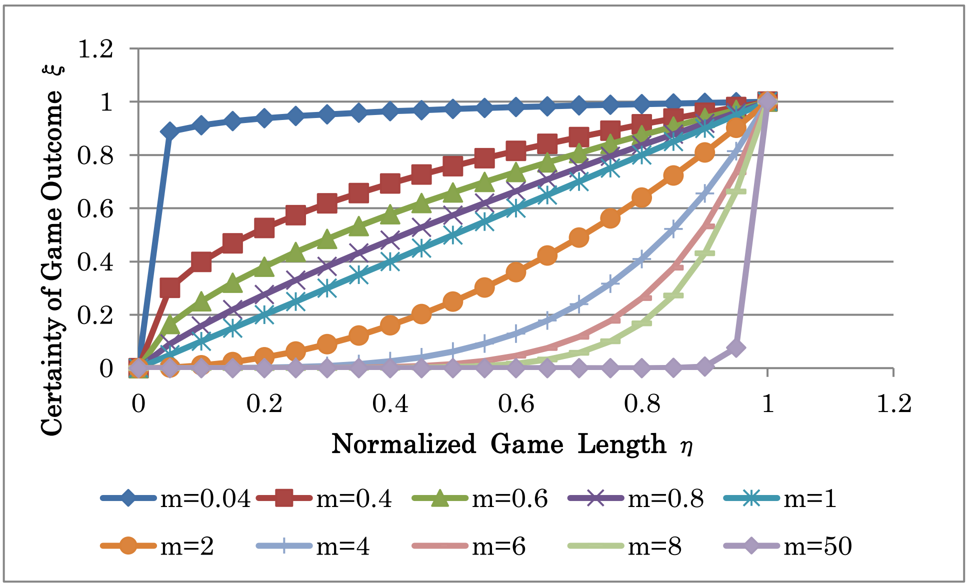

Figure 9 Certainty of game outcome ξ against normalized game length η.