FACETS OF UNCERTAINTY
Kos Island, Greece | 17–19 October 2013
European Geosciences Union
International Association of Hydrological Sciences
International Union of Geodesy and Geophysics
5th EGU LEONARDO CONFERENCE • HYDROFRACTALS '13 • STATISTICAL HYDROLOGY—STAHY '13

Windows of predictability in dice motion

Panayiotis Dimitriadis, Katerina Tzouka and Demetris Koutsoyiannis

Department of Water Resources and Environmental Engineering
Faculty of Civil Engineering
National Technical University of Athens, Greece

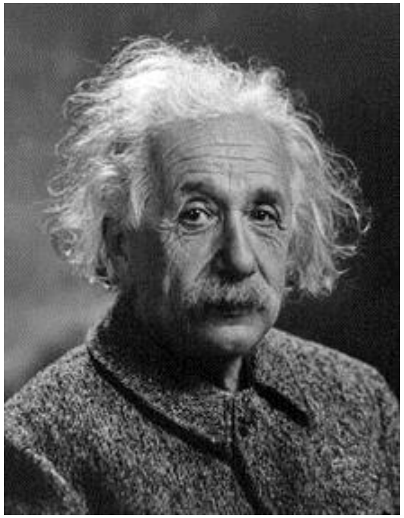
Presentation available online: http://itia.ntua.gr/1394/

Dice games are old

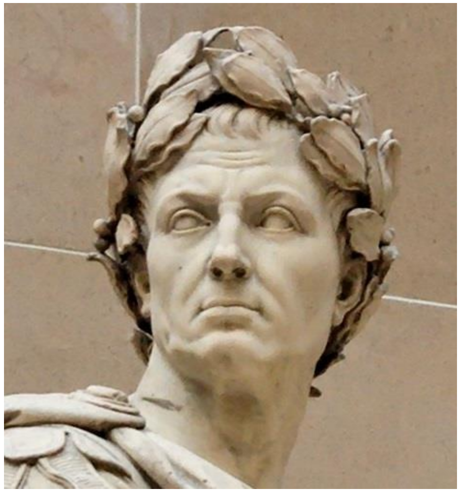
- All these dice are of the period 580-570 BC from Greek archaeological sites:
  - Left, Kerameikos Ancient Cemetery Museum, Athens, photo by D. Koutsoyiannis
  - Lower right and middle: Bronze die (1.6 cm), Greek National Archaeological Museum, www.namuseum.gr/object-month/2011/apr/apr11-gr.html
  - Upper right: Terracotta die (4 cm) from Sounion, Greek National Archaeological Museum, www.namuseum.gr/object-month/2011/dec/dec11-gr.html
- Much older dice (up to 5000 years old) have been found in Asia (Iran, India).

Some famous quotations about dice

Αἰών παῖς ἐστι παίζων πεσσεύων
Time is a child playing, throwing dice
(Heraclitus; ca. 540-480 BC; Fragment 52)



Jedenfalls bin ich überzeugt, daß der nicht würfelt
I, at any rate, am convinced that He does not throw dice
(Albert Einstein, in a letter to Max Born in 1926):



Ἀνερρίφθω κύβος	***Iacta alea est***
Let the die be cast	The die has been cast
[Plutarch's version]	[Suetonius's version]

(Julius Caesar, 49 BC, when crossing Rubicon River)

Physical setting

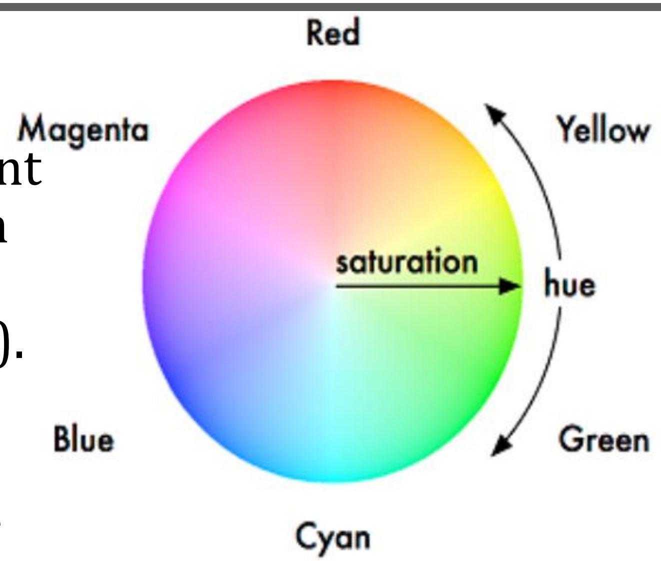
- The die motion is described by the laws of classical (Newtonian) mechanics and is determined by:
  - Die characteristics:
    - dimensions (incl. imperfections with respect to cubic shape),
    - density (incl. inhomogeneities).
  - Initial conditions that determine the die motion:
    - position,
    - velocity,
    - angular velocity.
  - External factors that influence the die motion:
    - acceleration due to gravity,
    - viscosity of the air,
    - friction factors of the table,
    - elasticity moduli of the dice and the table.
- Knowing all these, in principle we should be able to predict the motion and outcome solving the deterministic equations of motion.
- However the die has been the symbol of randomness (paradox?).

Some scientific studies on dice

- In a letter to Francis Galton (1894), W. F. Raphael Weldon, a British statistician and evolutionary biologist, reported the results of 26 306 rolls of 12 dice; the outcomes show a statistically significant bias toward fives and sixes (observed frequency 0.3377 against theoretical 0.3333; see Labby, 2009).
- Labby (2009) repeated Weldon's experiment (26 306 rolls of 12 dice) after automating it and reported outcomes close to those expected for fair dice (probabilities ~1/6, no autocorrelation).
- Strzalco et al. (2010) claim that a die is not fair by dynamics as the probability of the die landing on the face that is the lowest one at the beginning is larger than on the other faces.
- The same claim is made by Kapitaniak et al. (2012) who conclude that the die throw is neither random nor chaotic.
- Grabski et al. (2010) and Nagler and Richter (2008) call the dice behaviour pseudorandom because the motion is governed by deterministic laws (albeit with high sensitivity to initial conditions).

Researchers and apparatus for the experiment

Technical details



Color wheel of primary colours hue and saturation (www.highend.com/support/controllers/documents/html/en/sect-colour_matching.htm)

- Each side of the die is painted with a different colour: blue, magenta, red, yellow and green (basic primary colours) and black (highly traceable from the video as the box is white).
- The visualization is done via a camera with frame frequency of 120 Hz. The video is analyzed to frames and numerical codes are assigned to coloured pixels (based on the HSL system) and position in the box (two Cartesian coordinates).
- The area of each colour traced by the camera is estimated and then non-dimensionalized with the total traced area of the die. Pixels not assigned to any colour (due to low camera analysis and blurriness) are typically ~30% of the total traced die area.
- In this way, the orientation of the die in each frame is known (with some observation error) through the colours shown looking from above.
- The audio is transformed to a non-dimensional index from 0 to 1 (with 1 indicating the highest noise produced in each video) and can be used to locate the times in which the die hits the bottom or sides of the box.

Experiments made

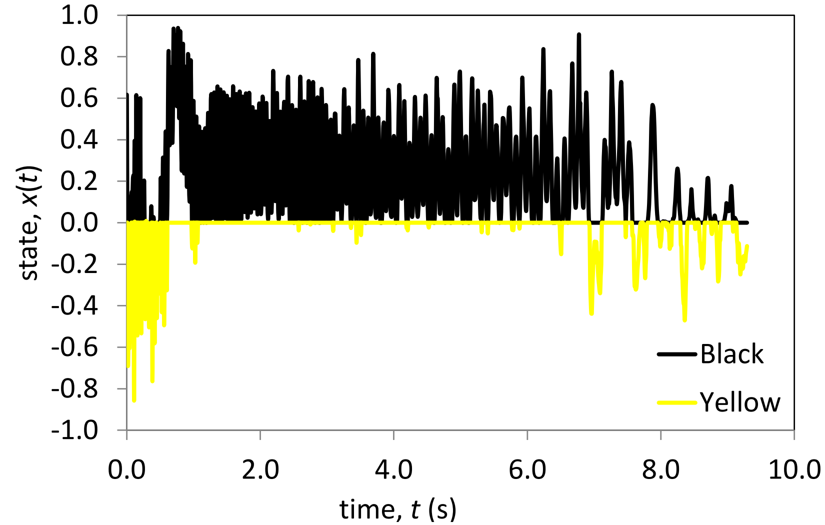
- In total, 123 die throws were performed, 52 with initial angular momentum and 71 without.
- The height from which the die was thrown remained constant for all experiments (15 cm).
- However, the initial orientation of the die varied .
- The duration of each throw varied from 1 to 9 s.

A selection of frames from die throws 48 (upper left) and 78 (lower left) and video for 78 (right).

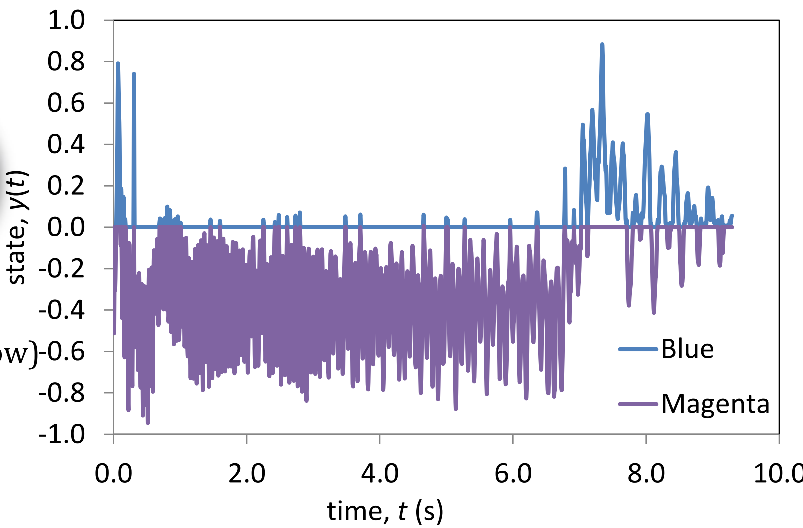
Representation of die orientation

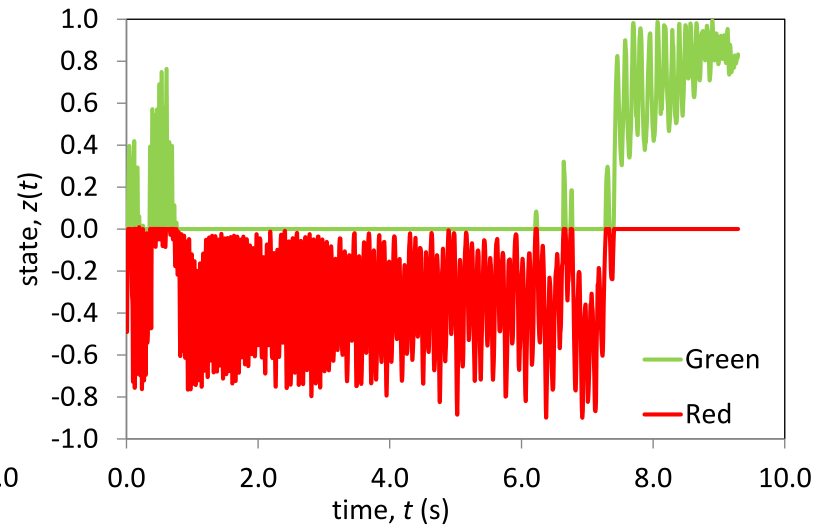
- The evolution of die orientation is most important as it determines the outcome.
- The orientation can be described by three variables representing proportions of each colour, as shown from above, each of which varies in [−1,1] (see table and figures which show raw values for experiment 78).

Value →	−1		+1	
Variable ↓	Colour	Pips	Colour	Pips
x	yellow	1	black	6
y	magenta	3	blue	4
z	red	5	green	2



Example:
$x = -0.25$ (yellow)
$y = 0.4$ (blue)
$z = -0.35$ (red)

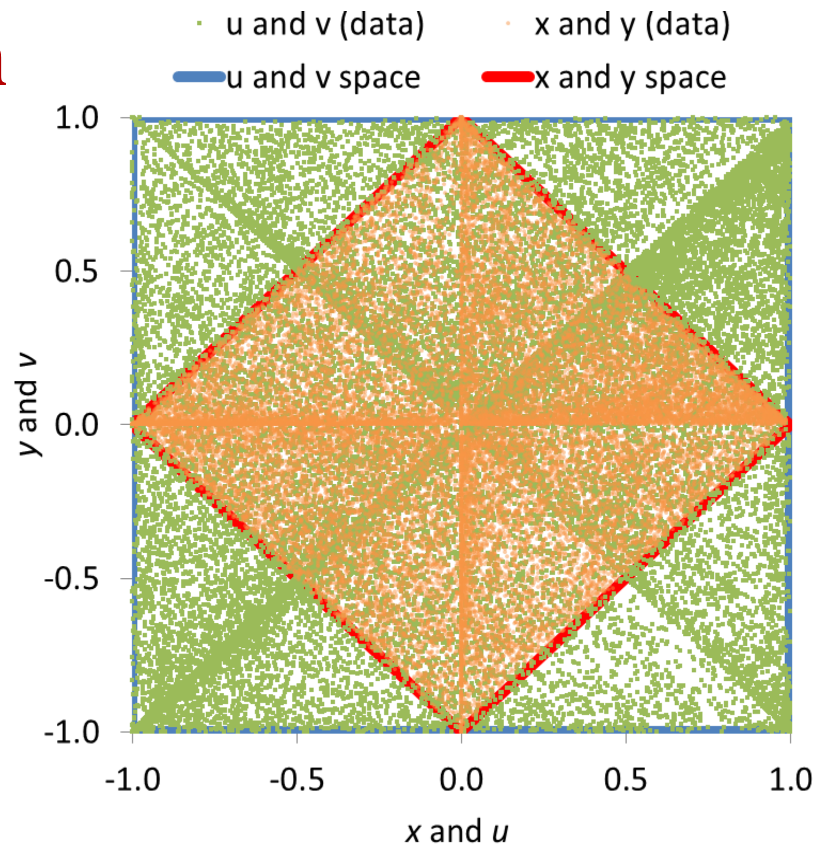


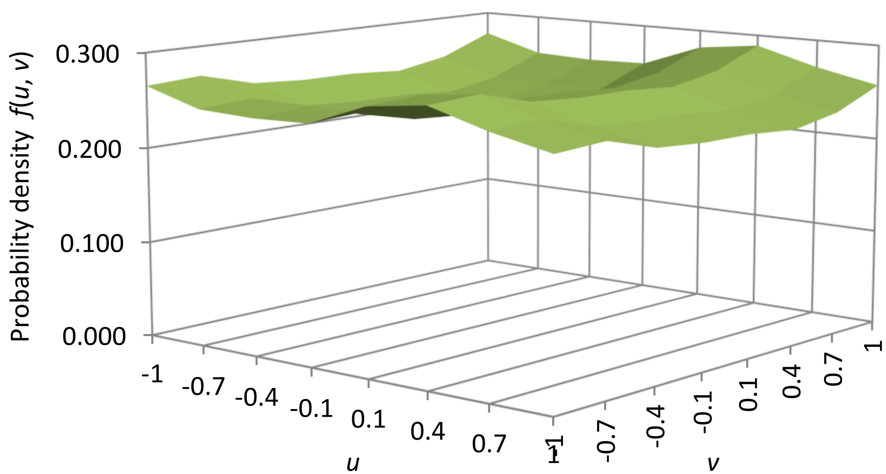


Alternative representation

- The variables x, y and z are not stochastically independent of each other because of the obvious relationship $|x| + |y| + |z| = 1$.
- The following transformation produces a set of independent variables u, v, w, where u, v vary in $[-1,1]$ and w is two-valued $(-1,1)$:

$$\begin{Bmatrix} u = x + y \\ v = x - y \\ w = \text{sign}(z) \end{Bmatrix} \leftrightarrow \begin{Bmatrix} x = (u + v) / 2 \\ y = (u - v) / 2 \\ z = w(1 - \max(|u|, |v|)) \end{Bmatrix}$$

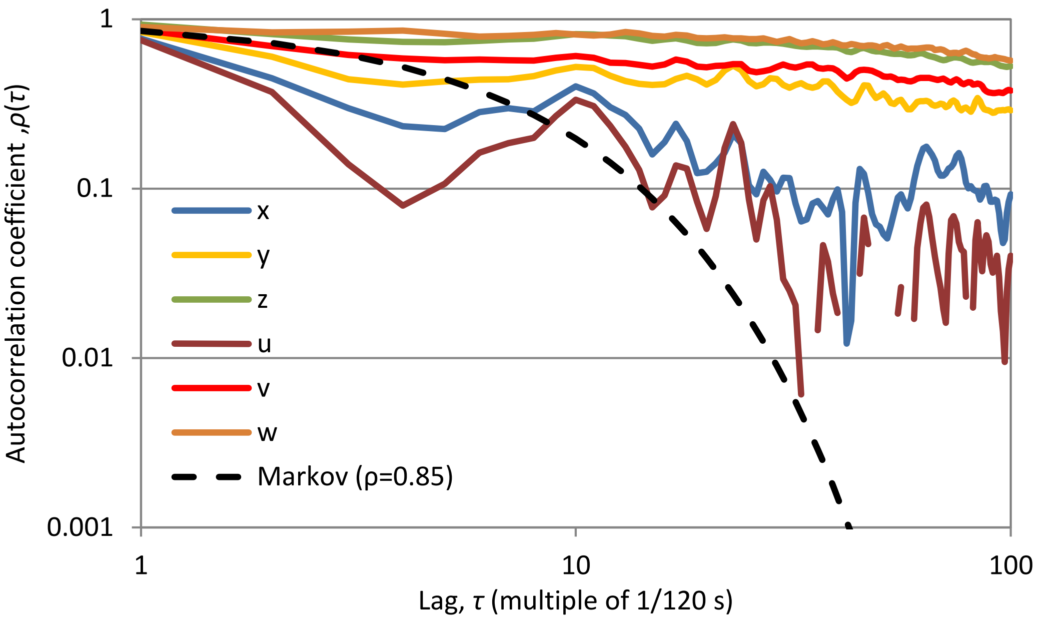


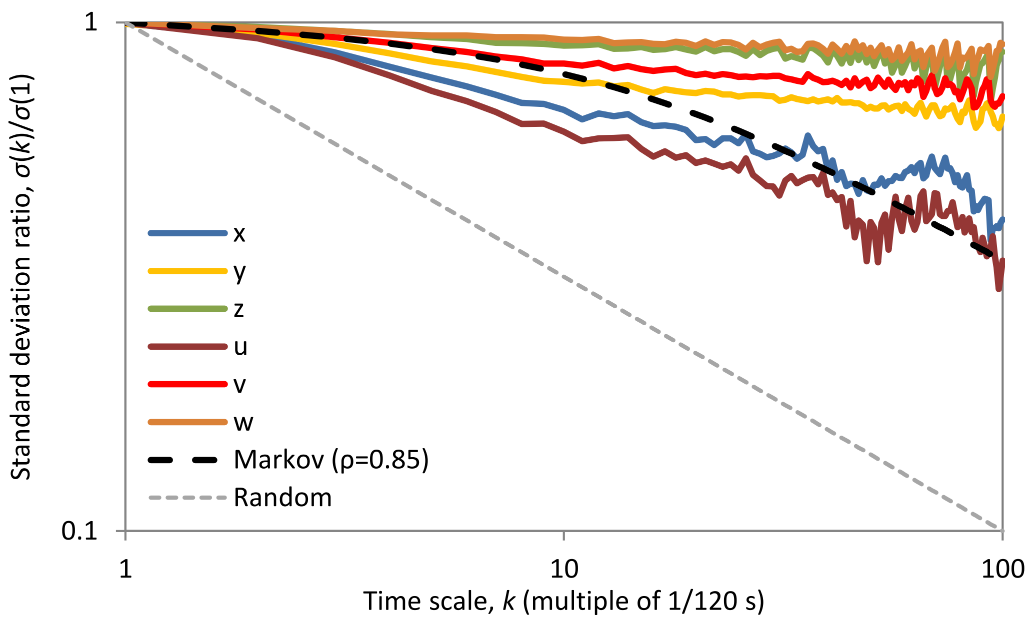


The plot of all experimental points and the probability density function show that u and v are independent and fairly uniformly distributed except that states for which $u \pm v = 0$ (corresponding to one of the final outcomes) are more probable.

Stochastic behaviour

- Autocorrelograms and climacograms (here those for experiment 78 are shown) indicate:
  - Strong dependence in time;
  - Long-term, rather than short-term persistence.
- Strong dependence enables stochastic predictability.





Stochastic models

- Two parsimonious (3-parameter) linear stochastic models were tested, which predict the state $\boldsymbol{s}((t+l)\Delta)$ based on a number of past states $\boldsymbol{s}((t-p)\Delta)$, where $p = 0, 1, \ldots$, $\Delta = 1/120$ s is the time step, $l\Delta$ is the lead time of prediction, and $\boldsymbol{s}$ denotes either the vector (x, y, z) or (u, v, w).
- **Model 1** (using the u-v-w formalism):
  $u((t+l)\Delta) = \sum_{p=1}^{10} a_p u((t-p+1)\Delta)$, $v((t+l)\Delta) = \sum_{p=1}^{10} a_p v((t-p+1)\Delta)$, $w((t+l)\Delta) = w(t\Delta)$
  where to reduce the number of parameters it was set $a_2 = a_3 = \ldots = a_9$.
- **Model 2** (using the x-y-z formalism):
  $\tilde{x}((t+l)\Delta) = \sum_{p=1}^{10} b_p x((t-p+1)\Delta)$, $\tilde{y}((t+l)\Delta) = \sum_{p=1}^{10} b_p y((t-p+1)\Delta)$,
  $\tilde{z}((t+l)\Delta) = \sum_{p=1}^{10} b_p z((t-p+1)\Delta)$
  where $b_2 = \ldots = b_9$. This is followed by adjustment to ensure consistency:
  $x((t+l)\Delta) = \tilde{x}((t+l)\Delta)/\tilde{s}$, $y((t+l)\Delta) = \tilde{y}((t+l)\Delta)/\tilde{s}$, $z((t+l)\Delta) = \tilde{z}((t+l)\Delta)/\tilde{s}$,
  where $\tilde{s} \coloneqq |\tilde{x}((t+l)\Delta)| + |\tilde{y}((t+l)\Delta)| + |\tilde{z}((t+l)\Delta)|$.
- The sets of parameters (a_1, a_2, a_{10}) and (b_1, b_2, b_{10}) depend on the lead time $l\Delta$. For each lead time they were numerically determined so as to minimize the mean square error over all time steps.
- To find an upper limit for predictability the entire data set of an experiment was used (no model validation period).

Deterministic data-driven model

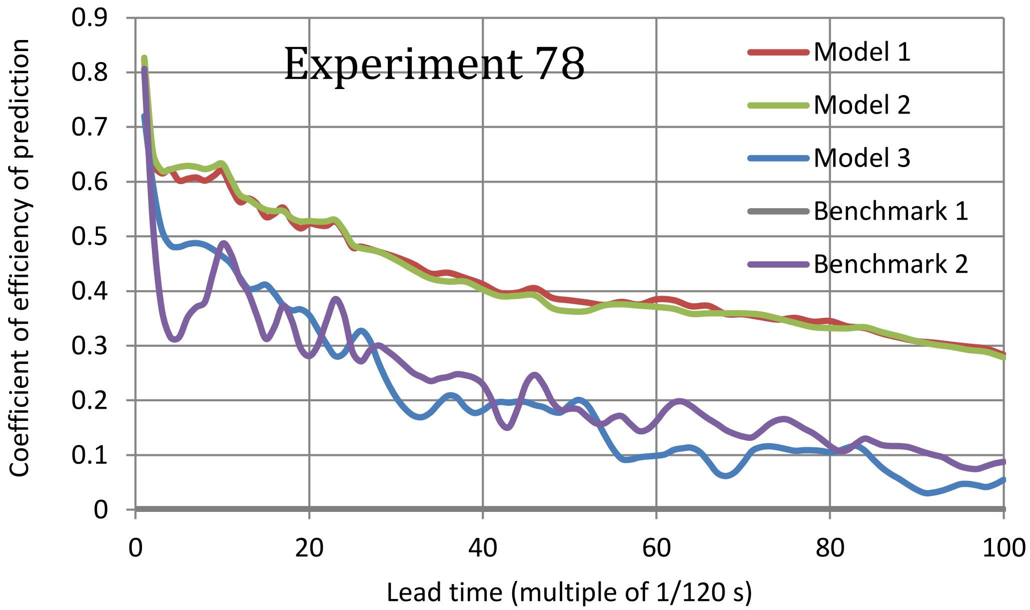
- **Model 3** is a deterministic model, purely data-driven, known as the analogue model (e.g. see Koutsoyiannis et al. 2008); it does not use any mathematical expression between variables.
- To predict $\boldsymbol{s}((t+l)\Delta)$, based on past states $\boldsymbol{s}((t-p)\Delta)$, $p = 0, 1, \ldots, m$, where $\boldsymbol{s} = (x, y, z)$:
  - We search the data base of all experiments to find similar states (neighbours or analogues) $\boldsymbol{s}^i((t^i - p)\Delta)$, so that $\sum_{p=1}^{m} \left\| \boldsymbol{s}^i\left((t^i - p)\Delta\right) - \boldsymbol{s}\left((t-p)\Delta\right) \right\|^2 \leq c$, where c is an error threshold.
  - Assuming that n such neighbours are found, for each one we find the state at time $(t^i + l)\Delta$, i.e. $\boldsymbol{s}^i((t^i + l)\Delta)$ and calculate an average state $\tilde{\boldsymbol{s}}((t+l)\Delta) = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{s}^i\left((t^i + l)\Delta\right)$.
  - We adjust $\tilde{\boldsymbol{s}}((t+l)\Delta)$ to ensure consistency, in the same manner as in Model 2.
- After preliminary investigation, it was found that a number of past values $m = 10$ and a threshold $c = 0.5$ work relatively well.

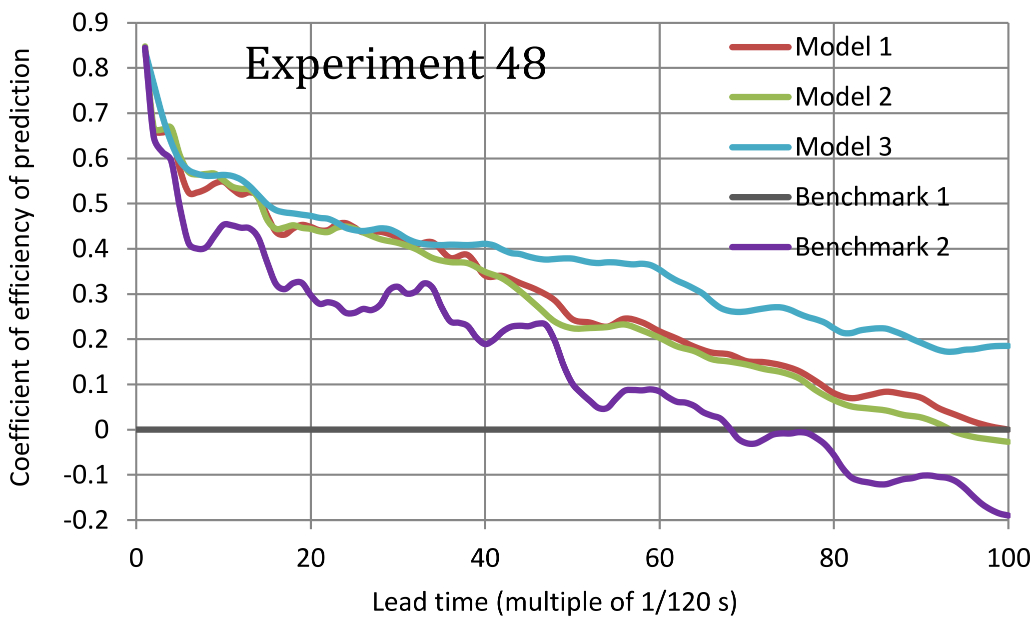
Benchmark models

- The three prediction models are checked against two naïve benchmark models.
- In **Benchmark 1** the prediction is the average state, i.e. $\boldsymbol{s}((t+l)\Delta) = \mathbf{0}$. Although the zero state is not permissible per se, the Benchmark 1 is useful, as any model worse than that is totally useless.
- In **Benchmark 2** the prediction is the current state, i.e. $\boldsymbol{s}((t+l)\Delta) = \boldsymbol{s}(t\Delta)$, regardless of how long the lead time $l\Delta$ is. Because of the high autocorrelation, it is expected that the Benchmark 2 will work well, for relatively small lead times.

Results

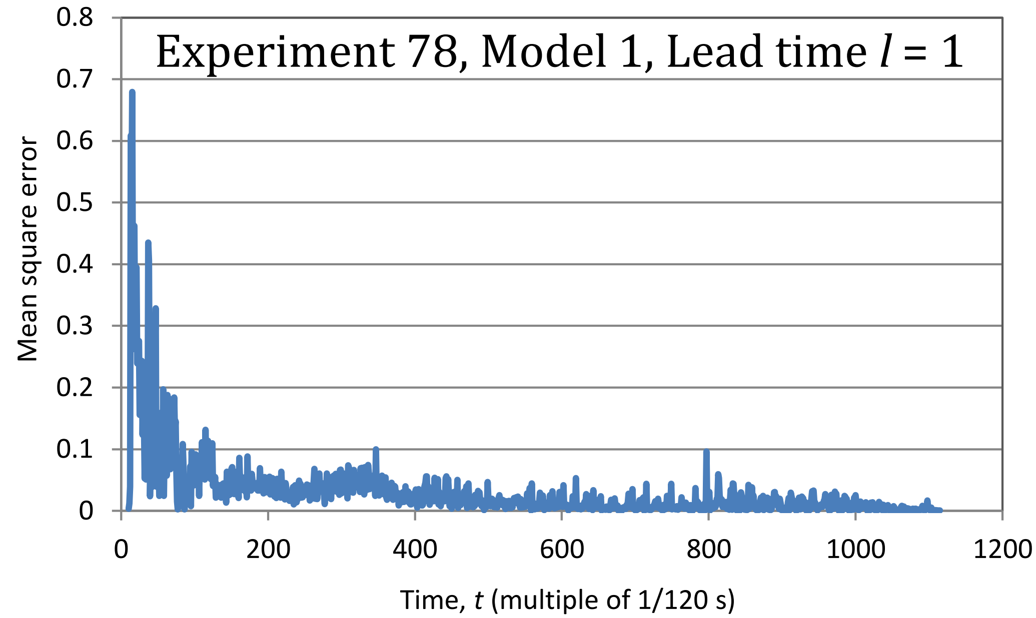
- For lead times $l\Delta \lesssim 1/10$ s, all three models, as well as Benchmark 2 provide relatively good predictions (efficiency $\gtrsim 0.5$)
- Predictability is generally superior than pure statistical (Benchmark 1) for lead times $l\Delta \lesssim 1$ s.
- Models 1 and 2 are virtually equivalent.
- Model 3 can be better or worse than models 1 and 2.

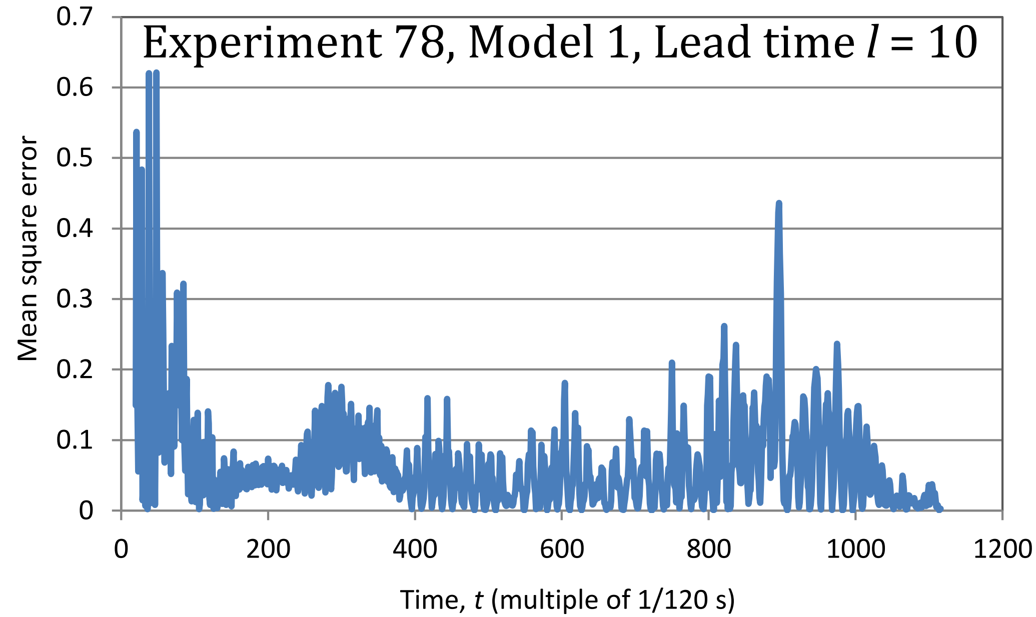




On the stationarity of the error

- Clearly, in increasing time, as the energy of the die dissipates, the error decreases and the predictability improves.
- The improvement of predictability is spectacular for small lead time (naturally, the error for the next frame tends to zero before the die stops).
- The situation worsens for larger lead times.





Concluding remarks

- There is no virus of randomness that affects dice.
- *Random* means none other than *unpredictable* or *unknown*.
- Both randomness and predictability coexist and are intrinsic to natural systems including dice (see Koutsoyiannis, 2010).
- Dice motion is both *deterministic* chaotic and *random*.
- Dice uncertainty is both *aleatory* (alea = dice) and *epistemic* (as in principle we could know perfectly the initial conditions and the equations of motion but in practice we do not).
- Dichotomies such as *deterministic* vs. *random* and *aleatory* vs. *epistemic* are false dichotomies.
- Dice behave like any other common physical system: predictable for short horizons, unpredictable for long horizons.
- The difference of dice from other common physical systems is that they *enable unpredictability* very quickly, at times < 1 s.

References

- Grabski, J., J. Strzalko and T. Kapitaniak, Dice throw dynamics including bouncing, *XXIV Symposium: Vibrations in Physicals Systems,* Vol. 24, Poznan-Bedlewo, 2010.
- Kapitaniak, M., J. Strzalko, J. Grabski and T. Kapitaniak, The three-dimensional dynamics of dice throw, *Chaos*, 22, 047504, 10.1063/1.4746038, 2012.
- Koutsoyiannis, D., H. Yao and A. Georgakakos, Medium-range flow prediction for the Nile: a comparison of stochastic and deterministic methods, *Hydrological Sciences Journal*, 53 (1), 142–164, 2008.
- Koutsoyiannis, D., A random walk on water, *Hydrology and Earth System Sciences*, 14, 585–601, 2010.
- Labby, Z., Weldon’s dice automated, *Chance*, 22 (4), 6-13, 2009.
- Nagler, J., and P. Richter, How random is dice tossing, *Physics Rev. E*, 78, 036207, 2008.
- Strzalko, J., J. Grabski, A. Stefanski and T. Kapitaniak, Can the dice be fair by dynamics?, *International Journal of Bifurcation and Chaos*, 20 (4), 1175-1184, 2010.