

Non-linear Dynamic Invariants Based on Embedding Reconstruction of Systems for Pedaling Motion

Juan-Carlos Quintana-Duque

Universität Konstanz

Introduction

Recent empirical studies of human motion (e.g. walking) have revealed complex dynamical structures, even under constant environmental conditions and several years of practice. These structures, also known as variability, have been used to determine disease severity, medication utility, and fall risk (Harbourne & Stergiou, 2009).

The term variability has been associated with different mathematical definitions related to chaotic systems, which are known as dynamic invariants. They are quantities describing the behavior of a dynamical system with the special property that the value of that quantity does not depend on the coordinate system.

If a dynamical system after a transient time converges to an attractor, i.e. a subset of an M -dimensional state space to which a dynamical system evolves, Takens' theorem says that it is possible to reconstruct the topological properties of this M -dimensional attractor using a projection in a N -dimensional embedding space created from observations of the system. That means the dynamic invariants can be calculated either from the M -dimensional state space or from a reconstructed N -dimensional embedding space.

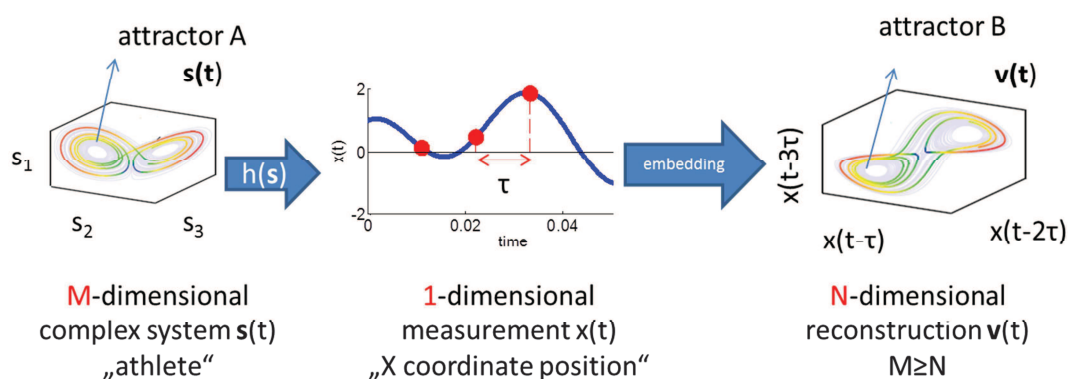


Figure 1. Embedding reconstruction of a multidimensional attractor

In dynamical systems of the real world, usually we don't have access to all the M -dimensional state space but we can do one-dimensional observations $h(s)$ of the system $s(t)$ (See Fig. 1). Takens' theorem guaranties that either observations of one state variable or observations of combination of state variables will be enough to reconstruct the attractor using an N -dimensional embedding space. For example,

we don't have access to all M -dimensional signals of the human body (e.g. neural and muscular signals) during pedaling but we can record the 3D spatial position of the knee joint (i.e. the position in a XYZ coordinate system). The time series data of the X coordinate of the knee joint position is enough to reconstruct the topological properties of the attractor using an N -dimensional embedding space provided such an attractor exists at all.

We found a lack of studies about the dynamic structure of pedaling motion and its potential benefits to distinguish subtle differences between pedaling motion patterns. In this paper, we evaluated two different methods for embedding (uniform and non-uniform) exploring their influence on two selected dynamic invariants based on embedding space (Maximal Lyapunov Exponent and Recurrence Period Density Entropy). The X coordinate of the knee joint motion was used as data which was captured during pedaling under low, medium and high workloads. Assuming that pedaling patterns change due to fatigue, our results show that it is possible to distinguish better between pedaling patterns (with and without fatigue) using dynamic invariants calculated from a reconstruction of the attractor in a non-uniform embedding instead of in a uniform embedding.

In the first part of this paper, we briefly review concepts of dynamical systems, embedding space, and the relation between human motion and dynamic systems. In the next section, we describe the experiment, the calculation of the embedding parameters and dynamic invariants. In the last section, we show the results and conclusions.

Concepts:

Dynamical systems and embedding space

A dynamical system consists of the phase space and the dynamics. The phase space is the collection of all possible world-states of the system in question. Each world state represents a complete snapshot of the system at some moment in time which is specified by a vector $x \in \mathbb{R}^M$ with M being the number of state variables. The dynamics can be described either by an M -dimensional map (discrete dynamics) or by a system of M first-order differential equations (continuous dynamics).

In continuous time, we don't specify the evolution of the state directly, but rather the rates of change of the components of the state $\frac{dx}{dt} = F(x, t)$. In the time-discrete case, the dynamics is an equation that transforms one point in the phase space (world-state), representing the state of the system "now", into another point (a second world-state), representing the state of the system one time unit "later". This equation is given by a function $x_{n+1} = f(x_n)$. Once an initial world-state is chosen (this is valid for both continuous and discrete time), the dynamics determine the world-state at all future times forming a trajectory in the phase space. Given a period of time long enough, the trajectory may eventually settle on a restricted geome-

try, called stable attractor, which is a portion of the state space, to which all trajectories converge asymptotically.

Usually, it is not possible to directly observe all components of the state variable $x \in \mathbb{R}^M$ of a high dimensional dynamical system but it is possible to observe the system obtaining 1D-measurements $y = h(x, \eta)$ providing a projection of the state space, where h is an observation function and η is an observational noise parameter (see Fig. 1). Given a sequence of state space points $x_n, n = 0, 1, \dots$, from the dynamics (discrete or continuous) of the underlying system, the sequence of discrete observations $y_n = h(x_n, \eta)$ can be used to create an N -dimensional embedding space. If the sequence of observations is long enough and contains low noise, the topological properties of an M -dimensional chaotic system can be reconstructed in the embedding space.

Embedding is the first step of non-linear time series analysis and is important because an adequate selection of the embedding parameters minimizes the effects of presence of noise and the finite amount of data. That means a good embedding can at the same improve the accuracy of the calculation of dynamic invariants (e.g. correlation dimension, Lyapunov exponents, and entropy) and reduce the prediction error in non-linear forecasting (Uzal, 2011).

There are two types of embeddings: Uniform and non-uniform (see Fig. 2). The classical uniform embedding is defined as $v_n = (y_n, y_{n-\tau}, \dots, y_{n-(d_e-1)\tau})$ which is determined by the embedding dimension d_e and embedding delay τ . The non-uniform embedding, which is a generalized version of uniform embedding, is determined by not equidistant consecutive delays $\tau_1 < \tau_2 < \dots < \tau_{d_e-1}$ and defined as $v_n = (y_n, y_{n-\tau_1}, y_{n-\tau_2}, \dots, y_{n-\tau_{d_e-1}})$.

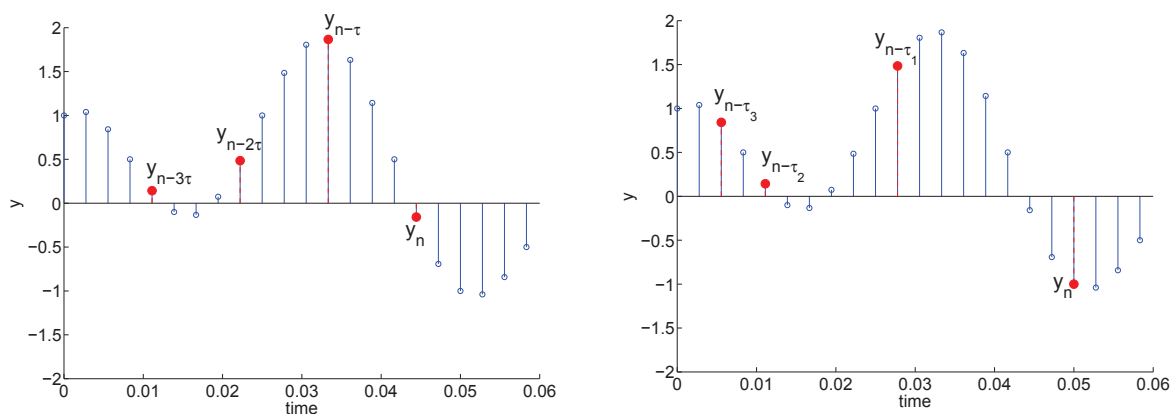


Figure 2. Examples of creation of four-dimensional embeddings from time series data. Uniform embedding (left) is defined with embedding delays $\tau < 2\tau < 3\tau$ as $v_n = (y_n, y_{n-\tau}, y_{n-2\tau}, y_{n-3\tau})$ whereas non-uniform embedding (right) is defined with embedding delays $\tau_1 < \tau_2 < \tau_3$ as $v_n = (y_n, y_{n-\tau_1}, y_{n-\tau_2}, y_{n-\tau_3})$

Takens' theorem gives the conditions under which a chaotic dynamical system can be reconstructed from a sequence of observations of the state of a dynamical system. The reconstruction preserves the properties of the dynamical system when embedding dimension d_e is large enough, i.e. $d_e \geq 2d + 1$, where d is the box-counting dimension of the attractor of the dynamical system. A precise knowledge of d_e is desirable since a large value of d_e will add redundancy and thus degrade the performance of many algorithms, such predictions and Lyapunov exponents (Kantz & Schreiber, 2004). However, in most of the cases we don't know the box-counting dimension d of the original attractor of real dynamical systems. Then, a suitable embedding dimension d_e must be estimated from the time series data y_n . On the other hand, this theorem doesn't say anything about how to select the consecutive delays $\tau_1, \tau_2, \dots, \tau_{d_e-1}$, which would be suited for computation. In previous works (for example, (Small, 2005) and (Stergiou, 2004)) a uniform embedding was assumed to be sufficient. However, recently it was shown that the estimates of topological invariants calculated from the numerical reconstruction of well-known attractors using non-uniform embedding were closer to theoretical values than using a uniform embedding, specially for quasiperiodic and multiple time-scale time series (Uzal, 2011).

Relation between human motion and dynamic systems

When the human body performs a cyclic motion, this tends to an attractive stable state, i.e. the motion trajectories show quasiperiodic but also fluctuations with erratic behavior. This is attributed to the body's ability to find the most stable solution coordinating all physiological systems over different timescales, whose behaviors are both highly variable and strongly dependent on each other and on environmental, bio-mechanical and morphological constraints. Furthermore, the motion is affected by a phenomenon called "next-cycle adaptation" which consists on modifying current acts based on perturbations occurring in antecedent ones.

A cyclical human motion can be modeled as a dynamic system (see Fig. 3), with a linkage system (body segments), actuators (muscles), sensors (e.g. tactile sensors, visual and vestibular system) and controller (Central Nervous System) and τ represents the time delay caused by transport and processing in the nervous system (Van der Kooij, 2008). In order to explain the "next-cycle adaptation" phenomenon, additional loops with large time delays can be included to expand the conceptual scheme showed in Fig. 3.

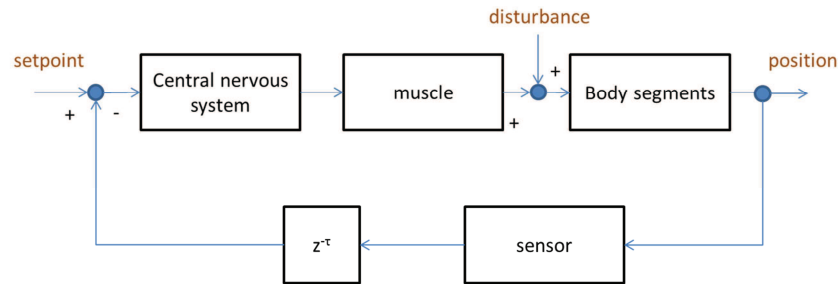


Figure 3. Conceptual scheme of the motion control of a musculoskeletal system (Van der Kooij, 2008)

These kinds of closed loop systems with time delays are described mathematically using delay differential equations. Buennen and colleagues in (Buennen, et al., 2000) gave a theoretical background for such types of systems. They suggested a special version of non-uniform embedding based on the fusion of two uniform embeddings to improve the analysis of dynamic properties. However, further analysis with other non-uniform embeddings was not done. Uzal in (Uzal, 2011) suggested that non-uniform embedding can also improve the analysis of systems with multiple periodicities which have similar properties as closed loop systems with time delays. These findings suggest that the dynamic properties of human motion data should be analyzed using a non-uniform embedding instead of a uniform embedding. Thus, it must be questioned whether the conclusions in previous works on analysis of gait motion and on physiological signals assuming only first order differential equations and a uniform embedding, are complete and correct.

Dynamic invariants: maximal Lyapunov exponent

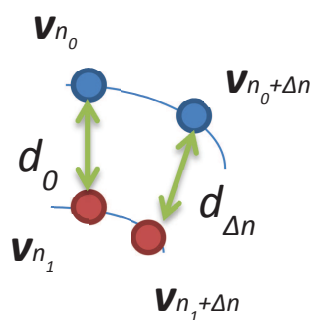


Figure 4. Graphical description of the procedure for estimating the maximal Lyapunov exponent. Defining $d_0 = \|v_{n_0} - v_{n_1}\|$ and $d_{\Delta n} = \|v_{n_0+\Delta n} - v_{n_1+\Delta n}\|$, this exponent is calculated from the logarithmic distance ratio $\frac{1}{\Delta n} \ln\left(\frac{d_{\Delta n}}{d_0}\right)$

Lyapunov exponents quantify the exponential divergence of initially close state-space trajectories over the course of time. When the attractor is chaotic, nearby trajectories on the attracting manifold diverge, on average, at an exponential positive rate characterized by the maximal Lyapunov exponent λ . Let v_{n_0} and v_{n_1} be two

points in the embedding space with Euclidian distance $\delta_0 = \|\mathbf{v}_{n_0} - \mathbf{v}_{n_1}\|$ and denote by $\delta_{\Delta n}$ the distance at some time Δn ahead between the trajectories emerging from these points, $\delta_{\Delta n} = \|\mathbf{v}_{n_0+\Delta n} - \mathbf{v}_{n_1+\Delta n}\|$ (see Fig. 4). Then, the maximal Lyapunov exponent λ is defined as $\delta_{\Delta n} = \delta_0 e^{\lambda \Delta n}$. If λ is positive, there is a strong signature of chaos in the time series data. The higher the instability and the divergence, the larger is the value of λ .

The maximal Lyapunov exponent λ is calculated numerically from a time series data as the slope of the average logarithmic divergence of the neighboring trajectories. These trajectories are described in the embedding space by l embedding points which are created from the time series data to be analyzed. The algorithm for calculating λ is the following: An embedding reference point \mathbf{v}_{n_0} is chosen and its nearest neighbors $\mathbf{v}_{n_1,2,\dots}$ with distance smaller than r are selected from the reference point's neighborhood \mathfrak{A} . Then, one computes the distances of all selected neighbors to the reference point following the trajectories as a function of the relative time Δn , (see Fig. 4). Repeating the latter for many values of n_0 and calculating the average of these results, i.e. $S(\Delta n) = \frac{1}{N} \sum_{n_0=1}^N \ln \left(\frac{1}{|\mathfrak{A}(\mathbf{v}_{n_0})|} \sum_{\mathbf{v}_n \in \mathfrak{A}(\mathbf{v}_{n_0})} \|\mathbf{v}_{n_0+\Delta n} - \mathbf{v}_{n+\Delta n}\| \right)$ where $1 < \Delta n < \Delta n_{max}$ and $N = l - \Delta n_{max}$, the influence of noise on the time series data is minimized and the fluctuations of the effective divergence will average out. If for some range of Δn the function $S(\Delta n)$ exhibits a robust linear increase, its slope is an estimated of the maximal Lyapunov exponent per time step. See more details in (Kantz & Schreiber, 2004).

Dynamic invariants: recurrence period density entropy (RPDE)

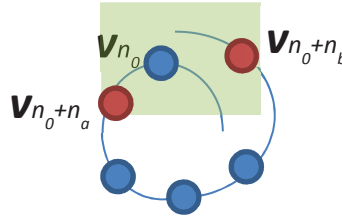


Figure 5. Graphical description of the procedure for estimating RPDE. The embedded point \mathbf{v}_{n_0} is followed forward in time observing the interceptions of its trajectory with the boundaries of a neighborhood placed around \mathbf{v}_{n_0} . The difference between the departure time n_a and the arrival time n_b is called the recurrence time.

RPDE characterizes the repetitiveness of a time series in phase space. A neighborhood \mathfrak{A} with size r is placed around the embedded data point \mathbf{v}_{n_0} . Then, the trajectory of the embedded point \mathbf{v}_{n_0} is followed forward in time, until it has left this neighborhood, (i.e. until $\|\mathbf{v}_{n_0} - \mathbf{v}_{n_0+n_a}\| > r$ for some $n_a > 0$, where $\|\cdot\|$ is the Euclidean distance) and continued further until the trajectory returns again to the neighborhood \mathfrak{A} (i.e. $\|\mathbf{v}_{n_0} - \mathbf{v}_{n_0+n_b}\| \leq r$ for $n_a < n_b$) (see Fig. 5).

The difference of these two discrete times is used to define the discrete recurrence time $T = n_b - n_a$. This procedure is repeated for all embedding reference points \mathbf{v}_{n_0} , forming a histogram of recurrence times $R(T)$ (Little, et al., 2007). This histogram is

normalized to give the recurrence time probability density $P(T) = \frac{R(T)}{\sum_{i=1}^{T_{\max}} R(i)}$, where T_{\max} is the maximum recurrence time found in the embedded space. The uncertainty of the recurrence time T is defined as the normalized entropy $RPDE = -\frac{\sum_{i=1}^{T_{\max}} P(i) \ln P(i)}{\ln T_{\max}}$. For a size r small enough, a value of 0 means that the data is perfectly periodic and a value of 1 means that it is a purely random signal.

Limitations of numerical calculation of dynamic invariants

Finite amounts of experimental data inevitably contain external noise due to environmental fluctuations and limited experimental resolution. In the limit of an infinite amount of noise-free data the numerical approaches for calculation of dynamic invariants (e.g. Lyapunov exponents) yield theoretical values of dynamic invariants. That means the accuracy of numerical estimation of invariants from experimental data depends on the quantity and quality of the data as well as on the quality of the reconstruction of the complexity of the dynamical system using an embedding space.

Methods

Data set

The pedaling motion of one healthy rider on a configurable static bike (Cyclus2) was recorded using a motion capture system (Lukotronic) with a sampling rate of 200 Hz. The average cadence was selected by the rider between 80 and 90 revolutions per minute (rpm) and the pedal break force (in Newton) was fixed with the bike simulator for each test according to the desired work load intensity. Three tests, each one of 6 minutes, were done with low (185 W with mean cadence of 80 rpm and pedal break force of 120 N), medium (210 W with 85 rpm and 130N), and high intensity power (240 W with 90 rpm and 140N) respectively with a short pause between them. The high power demand of the last test ensured that the rider started to be exhausted. Each test was divided in 3 intervals of 2 minutes and 1 minute was used from each interval as data for the calculation of invariants (see Fig. 6). The X coordinate of the knee motion was used as time series data. A Butterworth filter with a cut off frequency of 10 Hz was applied to remove measurement noise.

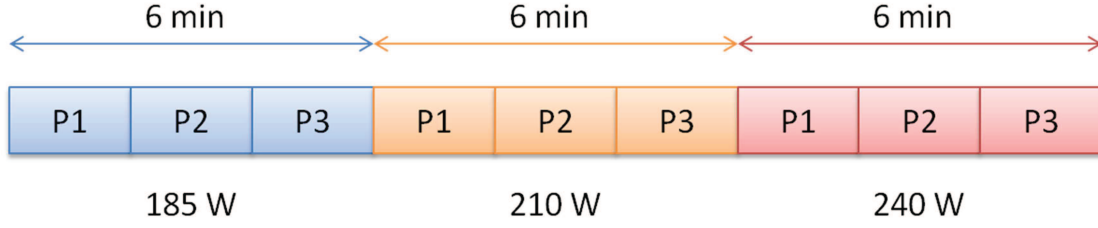


Figure 6. Interval schedule of the experiment

Calculation of the embedding parameters

For the uniform embedding $\mathbf{v}_n = (y_n, y_{n-\tau}, \dots, y_{n-(d_e-1)\tau})$, the embedding delay τ was selected from the first zero-crossing of the autocorrelation sequence of the time series y_n . The embedding dimension d_e was calculated using the false neighbor algorithm (FNN).

FNN consist on embedding a scalar time series y_n for a given embedding delay (in our case the previously calculated embedding delay τ) in increasingly higher dimensions. For each dimension, the Euclidean distance between pairs of embedding vectors \mathbf{v}_{n0} and their nearest neighbors \mathbf{v}_{n1} is calculated. If the current embedding with dimension k is sufficient to resolve the dynamics, an embedding with an extra dimension $k + 1$ will do it as well and the distance ratio between both embedding with k and $k + 1$ dimensions will not change. Then the dimension k is selected as embedding dimension d_e . See more details in (Kantz & Schreiber, 2004).

For the non-uniform embedding with not equidistant consecutive delays $\tau_1 < \tau_2 < \dots < \tau_{d_e-1}$, we use the method proposed by Uzal in (Uzal, 2011) which is based on the minimization of the average of a local cost function L over the attractor, in order to select the most appropriate parameters for the embedding reconstruction. The cost function L can be thought of as a sum of two terms with competing behavior $L = R + I$, where R should penalize redundancy when the window size τ_{d_e-1} is too small or, in case it is not, when the number of embedding components $d_e - 1$ within τ_{d_e-1} is insufficient to unfold the attractor. On the other hand, the I term should penalize irrelevance when the window size is too large or, in case it is not, when the number of components within τ_{d_e-1} is unnecessary larger than needed to unfold the attractor.

Uzal proposed in the definition of the cost function L to use the predictive power of a given reconstruction as a measure of redundancy R , and to use an irrelevance measure I based on the average distance between nearest neighbors in embedding space. The measure of redundancy has two free parameters: The number of nearest neighbors k and the prediction horizon T_M . A non-linear prediction is done for a given discrete prediction time T using a local constant model based on the first k nearest neighbors of reference point \mathbf{v}_{n0} . The mean of squared prediction error $E_k^2(T, \mathbf{v}_{n0})$ is computed for each embedding reference point \mathbf{v}_{n0} . Thus, the re-

dundancy value is the average of all prediction errors of all reference points over the discrete prediction time T on the interval $[0, T_M]$.

We assumed a four dimensional embedding with two embedding variables τ_1 and τ_2 defining the non-uniform embedding as $\mathbf{v}_n = (y_n, y_{n-\tau_1}, y_{n-\tau_2}, y_{n-(\tau_1+\tau_2)})$. We calculated the cost function of different combinations of τ_1 and τ_2 finding the optimal combination for each data set which minimizes this function (see Fig. 7).

Results and discussion

In Table 1, we show the embedding parameters for uniform and non-uniform embedding calculated from time series data of each interval (see Fig. 6).

	185W P1	185W P2	185W P3	210W P1	210W P2	210W P3	240W P1	240W P2	240W P3
d_e	4	4	4	4	4	4	4	4	5
τ	37	36	37	34	35	35	32	32	33
τ_1	4	12	3	13	11	13	13	11	4
τ_2	37	36	33	32	30	30	26	27	22

Table 1. Calculated embedding parameters for uniform (d_e, τ) and non-uniform (τ_1, τ_2) embedding

For the uniform embedding, the dimension d_e is the same in all intervals except for the time series of 240W P3. The embedding delay τ corresponds approximately to one quarter of the average pedal revolution in each interval. On the other hand, the non-uniform embedding parameter τ_1 seems more variable than τ_2 during each test. It suggests that the non-uniform embedding parameters change according to the dynamic properties of the time series to be analyzed.

In Figure 7, we show an example of the results of the calculation of the cost function proposed by Uzal (Uzal, 2011) applied on the time series of the interval 240W_P2 with the following parameters: Number of nearest neighbors $k = 2$ and prediction horizon $T_M = 30$. In this figure, we included a white line which corresponds to all possible values of τ_1 and τ_2 that yield equivalence to a uniform embedding (i.e. $\tau_2 = 2\tau_1$) and we plotted a point with coordinates $\tau_1 = 32$ and $\tau_2 = 64$ which represents the calculated τ for the uniform embedding (see interval 240W_P2 in Table 1). As we can see in this figure, the calculated uniform embedding doesn't correspond with optimal region colored with dark blue. It can be interpreted as the calculated uniform embedding, using as embedding delay the first zero-crossing of the autocorrelation sequence, doesn't provide the best reconstruction. A uniform embedding with small embedding delay (See in Fig. 7 the white line for small values of τ_1 and τ_2) gives a better reconstruction. However, the optimal reconstruction which minimizes the cost function is given by the non-uniform delays $\tau_1 = 11$ and $\tau_2 = 27$.

When the underlying dynamical system of the time series contains multiple

periodicities as the cyclical human motion seems to have (for example short- and long-term relations between motion cycles), uniform embedding may fail to capture all the periodicities and thus the embedding reconstruction becomes poor. Non-uniform embedding seems to take into account this characteristic.

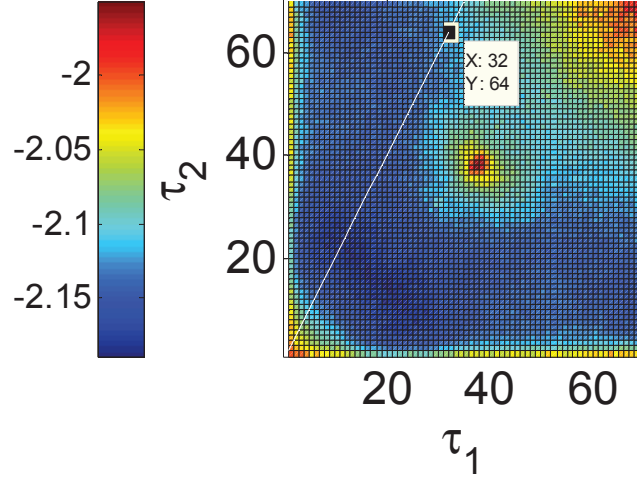


Figure 7. Selection of optimal embedding delays for non-uniform embedding using the Uzal's method. The optimal value is the lowest one. The white line represents the all possible values of τ_1 and τ_2 for uniform embedding.

The calculation of the maximal Lyapunov exponent is based on the slope where the average distance $S(\Delta n)$ increases linearly with Δn . However, $S(\Delta n)$ increases almost exponentially with appreciable oscillations when is calculated from the classical uniform embedding. Using a non-uniform embedding, the oscillations are almost imperceptible, making the calculation of this invariant more robust and reliable for pedaling data. (see Fig. 8)

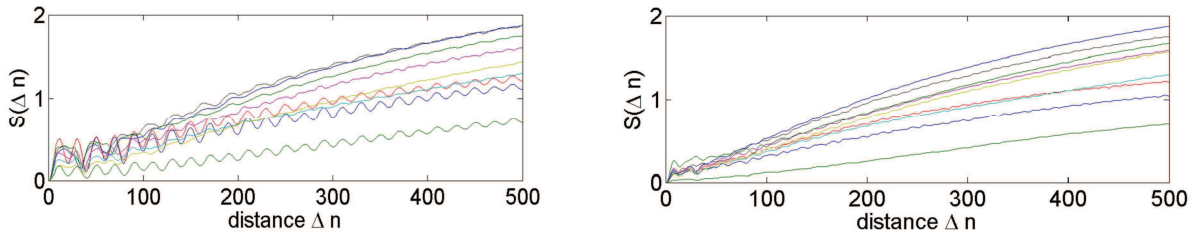


Figure 8. Curves for the calculation of maximal Lyapunov exponent using uniform (left) and non-uniform (right) embedding. Each curve represents a time series data.

Figure 9 plots the results of both invariants (maximal Lyapunov exponent and RPDE) applied on each of the nine time series obtained from the experiment, i.e. three workloads (185W, 210W and 240W), each one divided into three consecutive intervals (P1, P2, P3) (see Fig. 6). One expects in the plot of the calculated invariant values (see Fig. 9) to observe clear differences between the first and the last interval, i.e. between 185W_P1 and 240W_P3. Furthermore, one expects similar invariant values of intermediate intervals 185W_P2, 185W_P3, 210W_P1,

210W_P2 where the body had a stable motion, and an incremental tendency of the invariant values for the last intervals 210W_P3, 240W_P1, 240W_P2 and 240W_P3, where the pedaling motion was more irregular due to fatigue. These expected results are more clear using non-uniform embedding than using uniform embedding. The fatigue can be noticed especially in maximal Lyapunov exponent values but is not reflected in RPDE values. The workload intensity is not reflected in the calculated invariants. More fatigue tests with other riders, including physiological measurements, are needed to verify the conclusions. Future work includes the study of other invariants and its application for motion classification.

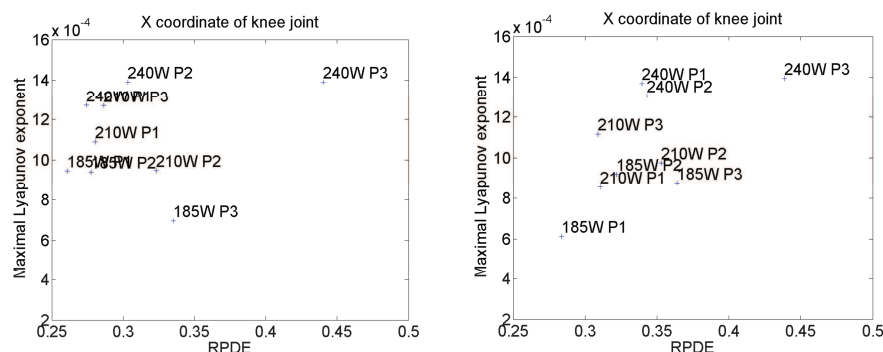


Figure 9. Invariant results using uniform (left) and non-uniform (right) embedding

Literature

- Bini, R., Diefenthaeler, F., & Mota, C. (2010). Fatigue effects on the coordinative pattern during cycling. *Journal of Electromyography and Kinesiology*, 20 (1), 102-107.
- Broer, H., & Takens, F. (2010). *Dynamical Systems and Chaos* (Bd. 172). Springer Verlag.
- Buenner, M., et al. (2000). Reconstruction of systems with delayed feedback: I. Theory. *The European Physical Journal D. Atomic, Molecular, Optical and Plasma Physics*, 10 (2), 165-176.
- Davids, K., Bennett, S., & Newell, K. (2006). Movement system variability.
- Harbourne, R., & Stergiou, N. (2009). Movement variability and the use of nonlinear tools: principles to guide physical therapist practice. *Physical therapy*, 89 (3), 267.
- Kantz, H., & Schreiber, T. (2004). *Nonlinear time series analysis* (Bd. 7). Cambridge Univ Press.
- Little, M., et.al. (2007). Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *BioMedical Engineering OnLine*, 6 (1), 23.
- Quintana, J.C. (2013). Evidence of Chaos in Indoor Pedaling Motion using Non-linear Methods. Accepted for publication. *Proceedings of Performance Analysis of Sport IX*. Routledge editorial.
- Small, M. (2005). *Applied nonlinear time series analysis: applications in physics, physiology and finance*. World Scientific Pub Co Inc.
- Sauer, T., Yorke, J. and Casdagli, M. (1991). Embedology. *Journal of Statistical Physics*. Springer, 65 (3).
- Stergiou, N. (2004). Innovative analyses of human movement. *Human Kinetics*.
- Uzal, L.C., Grinblat, G.L. and Verdes, P.F. (2011). Optimal reconstruction of dynamical systems: A noise amplification approach. *Physical Review E*, 84 (1).
- Van der Kooij, H., Koopman, B. and Van der Helm, F.C.T. (2008). *Human Motion Control*. Delft University.