ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-MOPOTK056

DATA-DRIVEN CHAOS INDICATOR FOR NONLINEAR DYNAMICS AND APPLICATIONS ON STORAGE RING LATTICE DESIGN*

Yongjun Li†, Robert Rainer, Brookhaven National Laboratory, Upton, New York, USA
Jinyu Wan[1], Yi Jiao[1], Institute of High Energy Physics, Beijing, China
Allen Liu, Dept. of Ele. Comp. Engr., Purdue University, W. Lafayette, Indiana, USA
[1]also at University of Chinese Academy of Sciences, Beijing, China

Abstract

A data-driven chaos indicator concept is introduced to characterize the degree of chaos for nonlinear dynamical systems. The indicator is represented by the prediction accuracy of surrogate models established purely from data. It provides a metric for the predictability of nonlinear motions in a given system. When using the indicator to implement a tune-scan for a quadratic Hénon map, the main resonances and their asymmetric stop-band widths can be identified. When applied to particle transportation in a storage ring, as particle motion becomes more chaotic, its surrogate model prediction accuracy decreases correspondingly. Therefore, the prediction accuracy, acting as a chaos indicator, can be used directly as the objective for nonlinear beam dynamics optimization. This method provides a different perspective on nonlinear beam dynamics and an efficient method for nonlinear lattice optimization. Applications in dynamic aperture optimization are demonstrated as real world examples.

INTRODUCTION

It is well-known that the predictability of motion in a nonlinear dynamical system is closely associated with its degree of chaos. Given an initial condition, although its motion is deterministic, its long-term prediction might not be quantitatively accurate because numerical errors can be cumulative and amplified. The Lyapunov exponent [1], i.e., the exponential growth of separation of infinitesimally close trajectories, is often used as a chaos indicator to characterize the sensitivity of chaotic motion to its initial condition.

Consider a different scenario: an unknown nonlinear dynamical system is encapsulated into a blackbox and only an ensemble of trajectories (input and output data) are available. Comparing actual trajectories to interpolated trajectories is one way to gauge chaos. A typical method to interpolate from known trajectories is to build a surrogate model with machine learning techniques. A surrogate model needs to be established first, then predictions can be made by evaluating trajectories with given initial conditions. This procedure is known as "supervised learning" [2]. To validate the model, the data is often randomly split into two clusters: a large training set and a small testing set. A model is then constructed from the training set. The performance of the model, i.e., the prediction accuracy, is measured by comparing the testing data against its prediction. The performance of the model depends on the type and complexity of the model, the volume of training data, the algorithm used for training, etc. Nevertheless, the prediction accuracy depends greatly on the degree of chaos. Therefore, an intuitive method for detecting chaos directly, purely from data is possible. In other words, predictability itself can act as a chaos indicator. From our studies we observed that by using the predictability of less-complex surrogate models, and a small volume of training data, some nonlinear behaviors in a dynamical system can be well characterized.

Surrogate models have been widely used in studying nonlinear dynamical systems [3–8], including charged particle motion in modern accelerators [9–15]. These models are obtained by training on either simulated data or experimental data, which have a high computational demand or require complicated experimental processing. If models can predict the dynamical system properties accurately with reduced resource requirements, they can be used for more efficient applications, such as optimization problems. Improving the prediction accuracy is the highest priority in these applications. In contrast to these existing approaches, the main advantage of using data-driven chaos indicators is that the requirement on the absolute accuracy of surrogate models is less demanding, and therefore can be structured with less complexity and data.

DATA-DRIVEN CHAOS INDICATOR FOR HÉNON MAP

The well-studied quadratic Hénon map, as shown in Eq. (1), is used as an example to demonstrate how to construct a data-driven chaos indicator for tune-scanning. It represents a thin sextupole kick followed by a linear rotation in a 2-dimensional phase space,

$$\begin{pmatrix} x \\ p \end{pmatrix}_{n+1} = \begin{pmatrix} \cos 2\pi\nu & \sin 2\pi\nu \\ -\sin 2\pi\nu & \cos 2\pi\nu \end{pmatrix} \begin{pmatrix} x \\ p - \lambda x^2 \end{pmatrix}_n , \quad (1)$$

where, n is a non-negative integer, ν is known as the linear tune of the transformation, and the sextupole strength λ is set as one for this demonstration. We assume the map in encapsulated as a blackbox with its tune as the control knob. For a given tune, some known trajectories that start with initial conditions (x_0, p_0) (input data) within a specific area, end with (x_n, p_n) (output data) after a limited number of turns. Based on the data, we can extract some parameters to characterize its long-term stability such as, the location of resonance lines and their stop-band widths, the relative size of the stable region, etc. This is accomplished by carrying

* Work mainly supported by US DOE Office of Science operated by BNL under Contract No. DE-SC0012704
† yli@bnl.gov

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-MOPOTK056

out a tune-scan. A tune-scan can be used to compare a nonlinear system's behavior at different linear tunes. At each given tune, some trajectories are produced from the blackbox. Most of them (around 80–90%) are used to train a surrogate model with a polynomial regression algorithm,

$$\begin{pmatrix} x \\ p \end{pmatrix}_n = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & a_4 \cdots \\ b_0 & b_1 & b_2 & b_3 & b_4 \cdots \end{pmatrix} \begin{pmatrix} 1 \\ x \\ p \\ x^2 \\ xp \\ p^2 \\ \vdots \end{pmatrix}_0 . \quad (2)$$

Here, the 7^{th} order polynomials are used. The rest, 10–20% of the data, is used as a testing set for performance validation. The validation is done by comparing the testing data against their model predictions (Fig. 1). Quantitatively, the prediction accuracy is measured with the mean squared errors (MSE) between the predictions and the true values. It also serves as the data-driven chaos indicator.

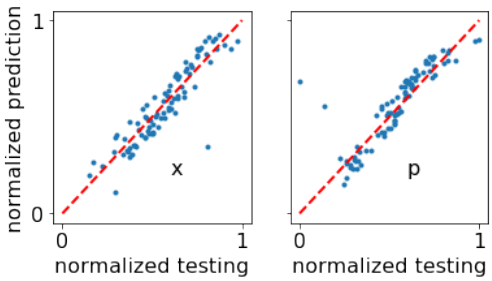

Figure 1: Prediction performance of 7^{th}-order polynomial surrogate model for the Hénon map.

The tune-scan result is illustrated in Fig. 2, in which the model's prediction performance is shown as the blue line with error bars. For each tune, the data was re-sampled randomly into different training/testing sets, multiple times (also known as the cross-validation technique). The shuffling of data can avoid selecting data that is trapped in a specific resonance, preventing the degree of chaos from being under- or over-estimated. The error bars represent the statistical fluctuations with different re-samplings. Due to quadratic perturbation, the worst model prediction occurs at $\nu = \frac{1}{3}$ as expected, which corresponds to a strong 3^{rd}-order resonance line. This resonance also has the widest stop-band width (approximated by the width of half-height of peak). Besides $\frac{1}{3}$, some other high order resonances at $\nu = \frac{1}{4}, \frac{1}{5}$, even $\frac{1}{7}$ are visible with this chaos indicator. For comparison, a long-term (2,048 turns) transformation starting from a wide initial condition of the x and p was computed. Its loss (i.e., unstable trajectory) rate as the function of the tune is also shown as the red solid line in Fig. 2. The data-driven chaos indicator observed appears to be highly correlated with the loss rate of long-term tracking.

It is interesting to note that an asymmetric stop-band width is detectable with this chaos indicator in Fig. 2. The appear-

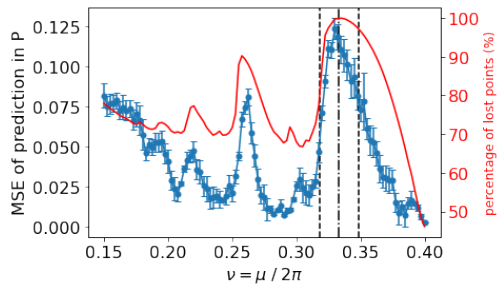

Figure 2: Prediction performance of a polynomial surrogate model (blue line with error bars) vs. loss rate (solid red) of Hénon map at different tunes.

ance of the asymmetry is due to amplitude dependent detuning. It behaves differently when the linear tune is slightly off the resonance line as shown in Fig. 3. When the linear tune is below the $(\frac{1}{3})^-$ resonance (in the left side), its amplitude-dependent tunes drift away from the resonance (dashed line). Therefore, the motions are less chaotic, and the left stop-band width is narrow. But when the linear tune is above the $(\frac{1}{3})^+$ (in the right side), its amplitude-dependent tune merges to the resonance quickly (solid line), so the motions are more chaotic, and the stopband width at the right side is correspondingly wide. This asymmetry is also observed at $\nu = \frac{1}{4}, \frac{1}{5}$ and confirmed with the loss rate.

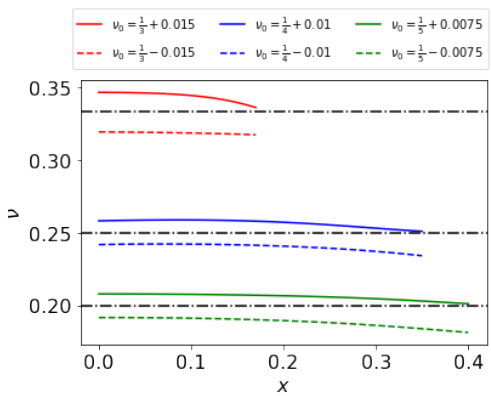

Figure 3: Different behaviors of the amplitude-dependent detuning when the linear tune approaches the $\frac{1}{3}, \frac{1}{4}, \frac{1}{5}$ resonances, which results in asymmetric stop-band widths.

The two tune-scan results in Fig. 2 are closely correlated. The information needed for these, however, can be very different. Using the data-driven chaos indicator, even short-term (20 turns) map transformations for only partial initial conditions can provide some useful information. The tune-scan using the loss rate is more accurate, but it also requires a greater number (2,048 turns) of map transformations for more initial conditions of x and p. In real-world applications, there may be a high resource demand to obtain such data. Using limited data resources to obtain an early chaos indicator has the potential advantage of boosting the optimization of design of a nonlinear dynamical system.

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-MOPOTK056

DATA-DRIVEN CHAOS INDICATOR FOR STORAGE RINGS

Consider a storage ring accelerator composed of various magnetic elements, in which the transportation of a charged particle for single turn (or a few repetitive turns) can be represented by a nonlinear transformation

$$\overrightarrow{X}_1 = M_{0\to1} \cdot \overrightarrow{X}_0. \quad (3)$$

Here, $\overrightarrow{X}_1, \overrightarrow{X}_0$ are the particle coordinates in the phase space, and $M_{0\to1}$ is the one-turn transportation map. Given the ring magnetic lattice, and using some simulated trajectory data $(\overrightarrow{X}_0, \overrightarrow{X}_1)$, a surrogate model can be constructed. The accuracy of the model, measured by the mean squared error (MSE) between the testing set and its model prediction, is a chaos indicator.

Using the existing NSLS-II lattice as an example, a surrogate model is used to approximate its one-turn transportation. The desired DA dimensions are $x = 25\ mm$ and $y = 10\ mm$ in the horizontal and vertical planes respectively, at the location of injection point. Thus, two elliptical areas in the phase space with axes at $(x, \frac{x}{\beta_x})$ and $(y, \frac{y}{\beta_y})$, are uniformly populated with 5,000 initial conditions as the input, $\overrightarrow{X}_0$. Here $\beta_{x,y}$ are the local Twiss parameters [16]. The one-turn transportation can be accomplished with particle tracking simulation. The coordinates at the exit are the output $\overrightarrow{X}_1$. The volume ratio of the training and testing data is 90%:10%. To avoid over- or under-fitting, the maximum number of training epochs was set to a sufficiently large number, and an early stopping point was used to halt the training once the model performance ceased improving. By varying harmonic sextupole settings, the accuracy of the model was monitored and used to drive a multi-objective genetic algorithm (MOGA) optimizer [17]. Besides using the four MSEs in each dimension (x, p_x, y, p_y) as the objectives, a minimum number of confined trajectories in the ensemble is used as a constraint. A good convergence of the average prediction accuracy was reached after a 100^{th} generation of MOGA evolution as shown in Fig. 4. Using 100 Intel® Xeon® 2.2-2.3 GHz CPU cores, optimization on this scale takes about 6 to 8 hours.

The candidate with the largest DA was chosen to implement a detailed frequency map analysis (FMA) [18] as shown in Fig. 5. An experimental test with live beam has been also carried out to confirm that this nonlinear lattice satisfies the requirements on the off-axis top-off injection and beam lifetime. Its DA is comparable to the solutions found using other methods [19–22].

The method was confirmed functional for the multi-bend achromat (MBA) type lattices, which are widely used in the diffraction-limited light source rings [23].

CONCLUSION

A novel data-driven chaos indicator concept was introduced by correlating the degree of chaos of a dynamical system and its surrogate model's prediction accuracy. This indicator can be used to optimize the dynamic aperture of storage rings. Traditionally, the prediction accuracy of a model has been critically important for many machine-learning applications. With this method, however, the prediction accuracy is used as a relative indicator of the chaos of a dynamical system. Greater accuracy is therefore less important, and surrogate models which have a lower resource demand are sufficient for this purpose. This method also provides a new perspective on the characterization of chaos in nonlinear dynamical systems and an efficient method for dynamic aperture optimization.

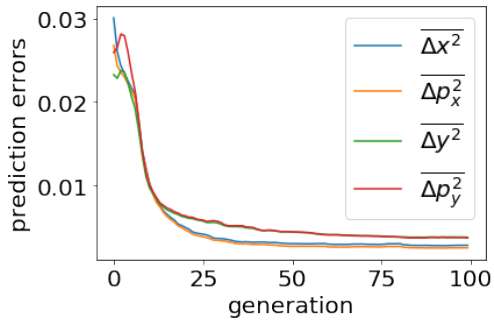

Figure 4: Convergence of prediction accuracy measured with four mean squared errors of test datasets in the surrogate/MSE MOGA optimization.

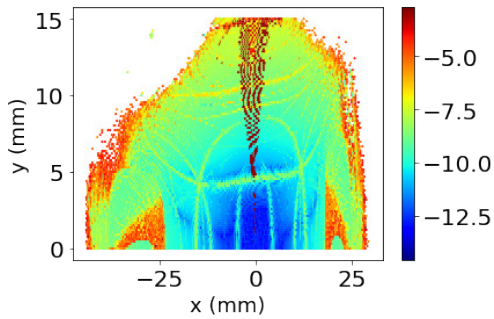

Figure 5: On-momentum dynamic aperture colored with the diffusion, obtained from frequency map analysis for the candidate with the largest DA. The color represents the tune diffusion $\log_{10}(\Delta\nu_x^2 + \Delta\nu_y^2)$ [18].

ACKNOWLEDGMENTS

We would like to thank Dr. Y. Hao (MSU) and the colleagues from NSLS-II for stimulating discussions and support.

REFERENCES

[1] A. Vulpiani, F. Cecconi, M. Cencini, "Chaos: from simple models to complex systems", Vol. 17, World Scientific, 2009

[2] T. Mitchell "Machine Learning", McGraw Hill, 1997

[3] M. Barahona, C. Poon, "Detection of nonlinear dynamics in short, noisy time series", Nature 381 (6579), 1996, pp. 215–217.

ISBN: 978-3-95450-227-1 ISSN: 2673-5490 doi:10.18429/JACoW-IPAC2022-MOPOTK056

[4] T. Schreiber, A. Schmitz, "Improved surrogate data for nonlinearity tests", Phys. Rev. Lett. 77 ,1996, pp. 635–638

[5] I. Tokuda, T. Miyano, K. Aihara, "Surrogate analysis for detecting nonlinear dynamics in normal vowels", The Journal of the Acoustical Society of America 110 (6) 2001, pp. 3207–3217

[6] A. Deshmukh, J. Allison, "Design of nonlinear dynamic systems using surrogate models of derivative functions", in: ASME 2013 International Design Engineering Technical Conferences and Computers and Information in 21 Engineering Conference, American Society of Mechanical Engineers Digital Collection, 2013

[7] K. Lindhorst, M. Haupt, P. Horst, "Efficient surrogate modelling of nonlinear aerodynamics in aerostructural coupling schemes", AIAA Journal 52 (9), 2014, pp. 1952–1966.

[8] C. Han, B. Glaz, M. Haile, Y. Lai, "Adaptable hamiltonian neural networks", arXiv preprint arXiv:2102.13235, 2021

[9] A. Sanchez-Gonzalez, *et al.*, "Accurate prediction of x-ray pulse properties from a free-electron laser using machine learning", Nature communications 8 (1), 2017, 1–9.

[10] F. Wang, *et al.*, "Machine learning for design optimization of storage ring nonlinear dynamics", arXiv preprint arXiv:1910.14220, 2019

[11] X. Huang, M. Song, Z. Zhang, "Multi-objective multi-generation gaussian process optimizer for design optimization", arXiv preprint arXiv:1907.00250, 2019

[12] A. Edelen, *et al.*, "Machine learning for orders of magnitude speedup in multiobjective optimization of particle accelerator systems", Phys. Rev. Accel. Beams 23, 2020, 044601

[13] J. Wan, P. Chu, Y. Jiao, "Neural network-based multiobjective optimization algorithm for nonlinear beam dynamics", Phys. Rev. Accel. Beams 23, 2020, 081601

[14] M. Kranjčević, *et al.*, "Multiobjective optimization of the dynamic aperture using surrogate models based on artificial neural networks", Phys. Rev. Accel. Beams 24, 2021, 014601

[15] J. Zhu,*et al.*, "High-fidelity prediction of megapixel longitudinal phase-space images of electron beams using encoder-decoder neural networks", Phys. Rev. Applied 16, 2021, 024005

[16] E. Courant, H. Snyder, "Theory of the alternating-gradient synchrotron", Annals of physics 3 (1), 1958, pp. 1–48.

[17] K. Deb, *et al.*, "A fast and elitist multi-objective genetic algorithm: NSGA-II", IEEE transactions on evolutionary computation 6 (2), 2002, pp. 182–197.

[18] J. Laskar, "Introduction to frequency map analysis, in: Hamiltonian systems with three or more degrees of freedom", Springer, 1999, pp. 134–150

[19] L. Yang, Y. Li, W. Guo, S. Krinsky, "Multiobjective optimization of dynamic aperture", Phys. Rev. ST - Accel. and Beams 14 (5), 2011, 054001

[20] Y. Li, L. Yang, "Multi-objective dynamic aperture optimization for storage rings", International Journal of Modern Physics A 31 (33), 2016, 1644019

[21] Y. Li, W. Cheng, L. Yu, R. Rainer, "Genetic algorithm enhanced by machine learning in dynamic aperture optimization", Physical Review Accelerators and Beams 21 (5), 2018, 054601

[22] Y. Li, *et al.*, "Fast dynamic aperture optimization with forward-reversal integration", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 988, 2021, 164936

[23] Y. Li, *et al.*, "Data-driven chaos indicator for nonlinear dynamics and applications on storage ring lattice design.", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 1024, 2022, 166060