ISBN: 978-3-95450-221-9 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2021-WEPV010

R&D OF THE KEK LINAC ACCELERATOR TUNING USING MACHINE LEARNING

A. Hisano†[1], M. Iwasaki#[1,2,3,4], I. Satake[5], M. Satoh[5,6]
H. Nagahara[4,3], N. Takemura[4,3], Y. Nakashima[4,3], and T. Nakano[3,4]
[1]Osaka City University Graduate School of Science, Osaka, Japan
[2]Nambu Yoichiro Institute of Theoretical and Experimental Physics (NITEP), Osaka, Japan
[3]Research Center for Nuclear Physics, Osaka University (RCNP), Osaka, Japan
[4]Osaka University Institute for Datability Science (IDS), Osaka, Japan
[5]High Energy Accelerator Research Organization (KEK), Ibaraki, Japan
[6]The Graduate University for Advanced Studies (SOKENDAI), Department of Accelerator Science, Kanagawa, Japan

Abstract

We have developed a machine-learning-based operation tuning scheme for the KEK e-/e+ injector linac (Linac), to improve the injection efficiency.

The tuning scheme is based on the various accelerator operation data (control parameters, monitoring data and environmental data) of Linac.

For the studies, we use the accumulated Linac operation data from 2018 to 2021.In this paper, we show the results on our R&Ds of, 1. visualization of the accelerator parameters (~1000) based on the dimensionality reduction, and, 2. accelerator tuning using the deep neural network (DNN). In the latter R&D, estimation of the accelerator parameters using DNN, and the accelerator simulator based on the Generative Adversarial Network (GAN) have been studied.

INTRODUCTION

We have developed an operation tuning scheme for the KEK e-/e+ injector linac (Linac) [1]. The Linac accelerator is a 600m long injector linac to distribute electrons and positrons to four ring accelerators: the Photon Factory (PF), the PF Advanced Ring (PF-AR), the SuperKEKB Electron Ring (HER) and the Positron Ring (LER). The Linac accelerator is capable of operating at up to 50 Hz in two bunches (96 ns apart), which is equipped with 100 beam position monitors (BPM), 30 steering magnets and 60 RF monitors.

Though the precise beam tuning and high injection efficiency are required, there exist several problems on the accelerator tuning as: 1. A lot of parameters (~1000) should be tuned, and these parameters are intricately correlated with each other; and 2. Continuous environmental change, due to temperature change, ground motion, tidal force, etc., affects to the operation tuning.

To solve the above problems, we have developed, 1. visualization of the accelerator parameters (~1000) trend/correlation distribution based on the dimensionality reduction to two parameters, and 2. accelerator tuning using the deep neural network, which is continuously updated with the accelerator data to adapt for continuous environmental change.to adapt the environment changes.

In this paper, we show the results on our R&D. For the studies, we use the electron beam data for Super KEKB injection accumulated in 2018 Nov. - 2021 Jun. This beam data includes the following parameters.

- 500 operating parameters (Steering magnet)
- 732 environmental parameters

In addition, we use the ratio of the upstream (SP_A1_M) and downstream (SP_58_0) charges of the accelerator as a quantitative measure of the injection efficiency of the accelerator.

VISUALIZATION OF THE ACCELERATOR PARAMETERS

In this study, we use dimensionality reduction with a Variational AutoEncoder (VAE) [2] as a visualization method of accelerator parameters. VAE is a neural network consisting of two networks: the encoder, which converts the input data into "latent variables" of arbitrary dimensions, and the decoder, which reconstructs the input data from the latent variables. Here, VAE assumes that "latent variables" are normally distributed, so that latent variables for similar trend input data are closely distributed when they are plotted on a space. Using this property, we can visualize the accelerator parameters by dimensionally reducing them to two-dimensional "latent variables" and plotting them in a two-dimensional space.

To check the visualisation performance by VAE, we created an accelerator parameters dataset containing 1232 parameters (operating params + env params). Figure 1 shows the results of the visualisation of the accelerator parameters accumulated between 2018 Nov. and 2021 Jun. using VAE trained by 2018 Nov to 2021 May data. The output results are coloured according to the injection efficiency of the input accelerator data.

As a result of Figure. 1, the accelerator parameter dataset containing 1232 parameters is visualized by VAE with di-

† m20sa029@uv.osaka-cu.ac.jp
# masako@osaka-cu.ac.jp

ISBN: 978-3-95450-221-9 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2021-WEPV010

mensionality reduction to two dimensions. And, the visualisation results show that in the short term(~1month) the accelerator trend does not drastically change, but in the long term(>3month) accelerator trend vary over a wide range.

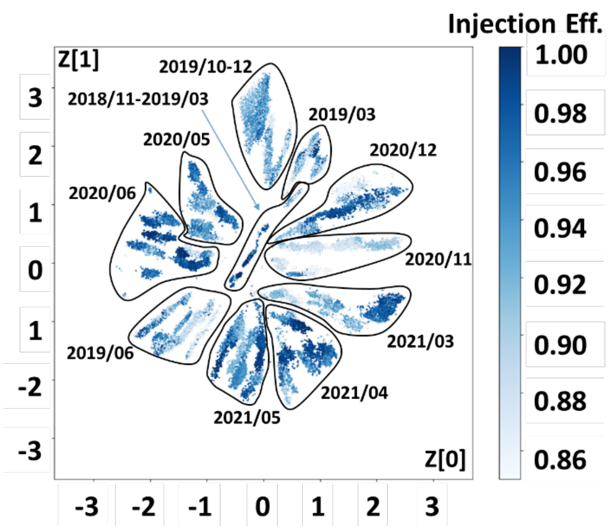

Figure 1: VAE output (two-dimensional Latent variables) from the 2018 Nov to 2021 June accelerator data. VAE is trained using 2018 Nov to 2021 May data. The output results are coloured according to the injection efficiency of the input accelerator data.

ACCELERATOR TUNING USING THE DEEP NEURAL NETWORK

The operating environment of the accelerator is continuously changing. In addition, the correlations between the parameters are complex. In this situation, to continuously adapt to the environmental changes to get the high injection efficiency, the environment driven DNN (Reinforcement ML) is a good candidate. However, to apply the Reinforcement ML to the operation tuning, we need 1. Estimation of the acc. parames using DNN, 2. realistic accelerator simulator. In the following, we introduce these R&D.

Estimation of the Acc. Parames using DNN

For the reinforcement learning, it is important to start from the optimized (the best) accelerator parameter value, so that ML doesn't need to search in wide parameter range.In addition, we need to obtain the relationship between efficiency and accelerator parameters, for the effective parameter optimization.

Therefore, we have developed a method to estimate an injection efficiency using DNN with accelerator parameters inputs, to assure to get the relationship between them.

In this study, we designed a Regression-DNN that takes 1232 accelerator parameters (operating params and env params) as input and outputs the injection efficiency. The implementation of the DNN is based on Tensorflow [3].

Evaluate the performance of DNN To evaluate the performance of DNN, we extract 150,000 events data from Linac data accumulated in 2018 Nov. - 2021 Jun. Note that the injection efficiency of the dataset is extracted to be uniformly distributed over 0.85. Out of 150,000 events, we use 135,000 events for training and 15,000 events for validation.

We trained and validated Regression-DNN on dataset described above. As a result, the mean squared error (MSE) between the predicted injection efficiency and the real injection efficiency is 0.00030 for the whole data interval. Figure. 2 shows injection efficiencies of predicted by DNN (orange) and real (blue), for 2020/11/30-2020/12/07 evaluation data. Vertical and horizontal axis indicate injection efficiency and date, respectively.

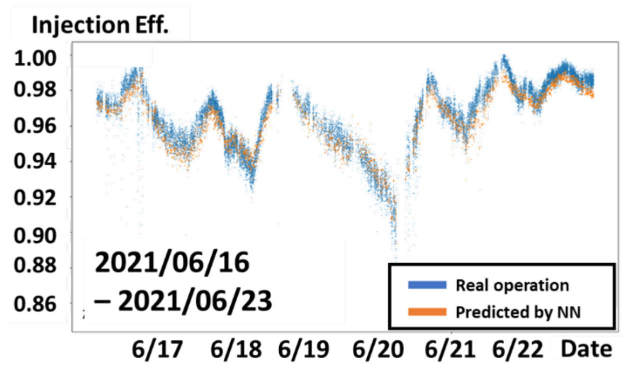

Figure 2: Injection efficiencies of predicted by DNN (orange) and real (blue), for 2020/11/30-2020/12/07 evaluation data. Vertical and horizontal axis indicate injection efficiency and date, respectively.

These results show that Regression-DNN can predict injection efficiency. In other words, Regression-DNN can describe the relationship accelerator parameters and injection efficiency.

Training DNN with past data Figure 1 shows that 2D latent value for 2021 Jun. is close to 2021 May. This indicates that the 2021 May accelerator data are similar to the 2021 Jun. accelerator data. So, using DNN trained with 2021 May data, we may be able to predict the injection efficiency in 2021 Jun. To test our hypotheses, we create three datasets with different inclusion periods. Figure 3 shows an overview of the three datasets.

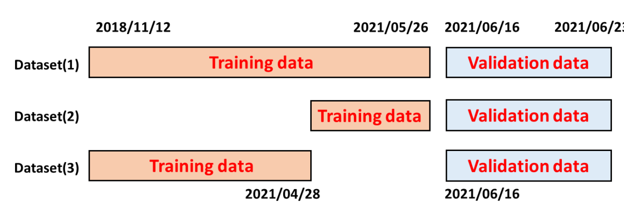

Figure 2: Overview of the datasets (1)-(3).

Each dataset contains 135,000 events for training and 15,000 events for validation. We trained and validated DNN on each dataset. As a result, the mean squared error of injection efficiency output by DNN trained on datasets (1)-(3) as validation data is 0.00038 (Dataset (1)), 0.00038 (Dataset (2)), 0.00300 (Dataset (3)). Figure 3 shows the comparison between the injection efficiency predicted by DNN (Dataset (1) orange, (2) red, (3) green) and the real

ISBN: 978-3-95450-221-9 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2021-WEPV010

incident efficiency (blue). Vertical and horizontal axis indicate injection efficiency and date, respectively.

The results show that the DNN trained with 2021 May data; dataset (1,2), can predict 2021 Jun. injection eff. In other words, in order to predict the injection efficiency with DNN trained on past data, it is necessary that the training data of DNN contains similar data to the trend of the period we want to predict.

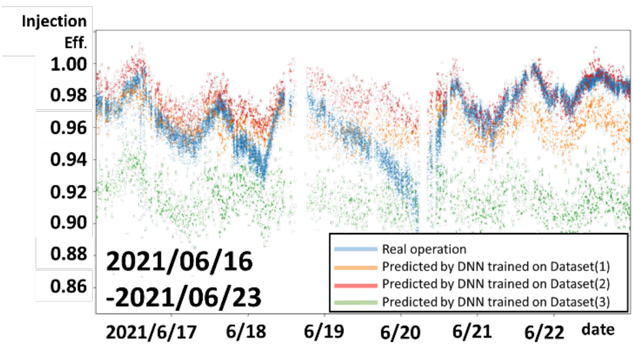

Figure 3: Predicted injection efficiencies for the 2021 June evaluation data, based on the DNN predictions with Dataset (1) to (3) (orange, red, and green), and the real operation injection eff.(blue). Vertical and horizontal axis indicate the injection eff. and date, respectively.

Accelerator Simulator Based on GAN

For the reinforcement learning, it is dangerous to do the parameter optimization based on the actual operation with the real accelerator. Therefore, realistic simulator is necessary for the pre-training of the actual reinforcement learning.

In this study, we develop a method using GAN (Generative Adversarial Network) [4]. Generative Adversarial Network (GAN) is a class of machine learning frameworks. Given a training set (real data), this technique learns to generate new data with the same statistics as the training set.

GAN consists of two networks (Generator and Discriminator) that compete with each other. The Generator is a network with noise data as input and new data with the same statistics as the training set as output. The Discriminator is a network with training set data(real) and output of Generator(fake) as input and discriminate between real and fake as output. By alternately training two competing networks, we can increase the similarity between the Generator output (fake data) and the training set (real data).

Using this property, by training GAN on the accelerator parameters data, it is expected to be possible to reproduce the data with keeping the correlation between the parameters. So, we verify whether the GAN trained by real data of the accelerator can generate data reproducing the correlation between parameters.

For GAN training, we create the following two datasets with different parameter configurations.

- Dataset(a) 2params
  Injection eff. + specific steering magnet param
- Dataset(b) 1233params
  Injection eff. + operating params+ env params

Please note that, Injection efficiencies in the datasets are uniformly extracted. And, each dataset contains 540,000 events for training and 60,000 events for validation.

Training by Dataset(a) We trained the GAN on the dataset (a), which consists of two parameters, the injection efficiency, extracted with a uniform distribution and specific steering magnet parameter (PY_32_4). Figure 4, shows a histogram of the magnet parameter (PY_32_4) reproduced by this trained GAN, and the real data used for training

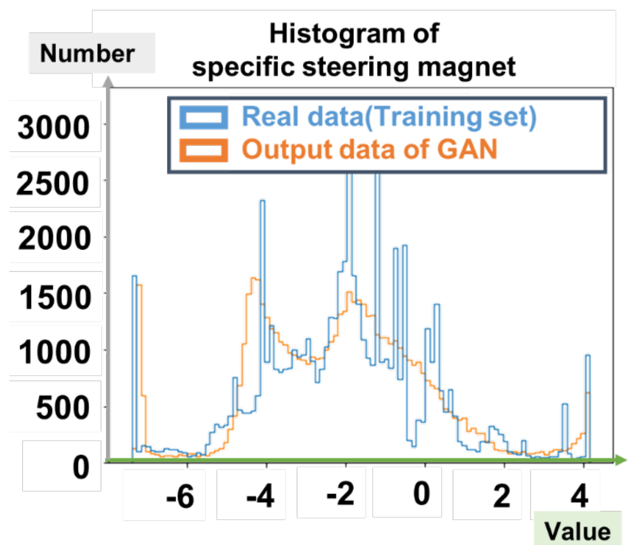

Figure 4: Histogram of the specific magnet (PY_32_4) param. reproduced by GAN trained Dataset(b) (orange) and the real data (blue).

In this case, from the results in Figure 4, we can see that GAN reproduces the characteristics (peaks, distribution) of real data by training. However, it is necessary to reproduce the correlation between the parameters as well as the characteristics of the individual parameters. Therefore, we examined the correlation between the injection efficiency and the steering magnet parameters in the dataset by making two-dimensional histograms. The created two-dimensional histogram is shown in Figure. 5.

ISBN: 978-3-95450-221-9 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2021-WEPV010

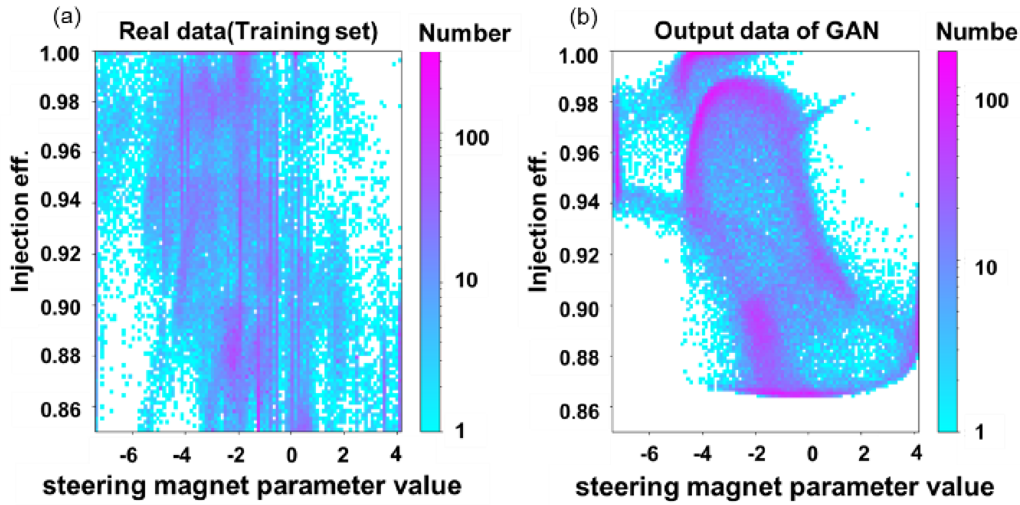

Figure 5: Two-dimensional histogram of real data(a) and reproduced data(b). Vertical and horizontal axis indicate Injection eff. and steering magnet parameter value.

From the results in Figure 5, we can see that GAN reproduces the two-parameter correlation trend. However, the GAN output data has peaks that did not match the real data and does not cover the full range of the real data.

Training by Dataset(b) We trained the GAN on the dataset (b), which consists of 1233 parameters, the injection efficiency extracted with a uniform distribution, steering magnet parameters and environmental parameters. Figure 6, shows a histogram of the magnet parameter (PY_32_4) reproduced by this trained GAN, and the real data used for training.

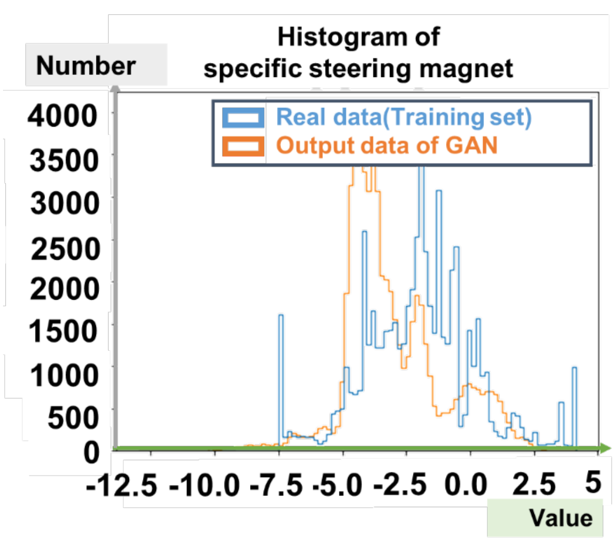

Figure 6: Histogram of the specific magnet (PY_32_4) param. reproduced by GAN trained Dataset(b) (orange) and the real data (blue).

From the results in Figure 6, we can see that GAN reproduces the characteristics (peaks, distribution) of real data by training. But we need to verify that the trained GAN is able to reproduce the correlations of the 1233 parameters, not just the individual parameters characteristics. Here, it is difficult to compare the real data with the reproduced data for the 1233 parameters. Therefore, we use the visualisation method with dimensionality reduction by VAE described above. Figure 7 shows two-dimensional histogram of the visualisation of the real data and reproduced data by GAN, using VAE trained by 2018 Nov to 2021 Jun data.

From the results in Figure 7, we can see that GAN output data dose not reproduce the whole period of the real data. Such a result, where the GAN can only reproduce part of the data used for training, is a possible event in the training of GAN, called "mode collapse".

SUMMARY

We have developed a machine-learning-based operation tuning scheme for the KEK e-/e+ injector linac (Linac), to improve the injection efficiency.

In this paper, we show the results on our R&Ds of, 1. visualization of the accelerator parameters (~1000) trend/correlation distribution based on the dimensionality reduction using VAE, 2. predicting injection efficiency using Regression-DNN, and 3. reproduction of accelerator parameters using GAN.

Using the electron beam data for Super KEKB injection accumulated in 2018 Nov. - 2021 Jun, we show that 1. It is possible to visualize the 1232 parameters by dimensionality reduction to two dimensions using VAE; 2. It is possible to predict the injection efficiency of accelerators using Regression-DNN. The prediction with the past data which has the similar parameter features, is effective to improve the prediction accuracy; 3. Accelerator simulator based on GAN reproduces the overall parameters correlation trend. On the other hand, GAN's output data dose not reproduce the whole period of the real data due to the "mode collapse", and further improvement is necessary to use for the pretraining of the reinforcement learning for the accelerator tuning system.

ISBN: 978-3-95450-221-9 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2021-WEPV010

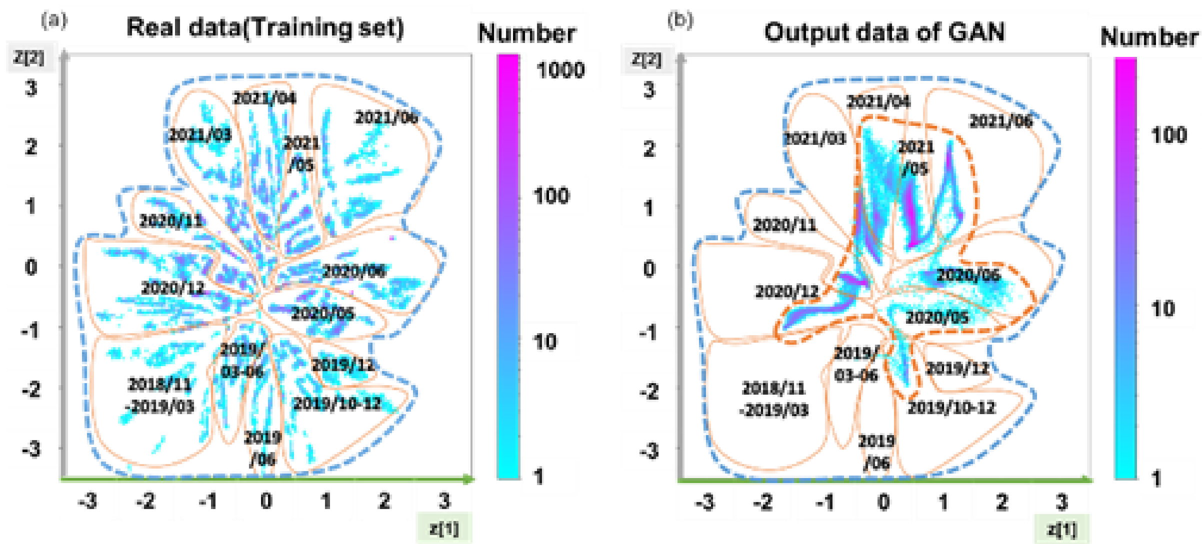

Figure 7: Two-dimensional histogram of real data(a) and reproduced data(b). Vertical and horizontal axis indicate VAE 2D latent value z [2] and z [1].

ACKNOWLEDGEMENTS

The R&D works are supported by the Research Project at the Research Center for Nuclear Physics (RCNP), Osaka University, and the Interdisciplinary Project of the Datability Frontier Organization, Osaka University. Part of this work is also supported by a collaborative research grant from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) under the "Diversity Research Environment Initiative (Traction Type)", and use of the Supercomputer System OCTOPUS at the Cybermedia Center, Osaka University.

REFERENCES

[1] Electron-positron injector (LINAC), https://www.kek.jp/en/Research/ACCL/LINAC/

[2] D. P. Kingma, and M. Welling, "Auto-Encoding Variational Bayes", arXiv:1312.6114, 2014.

[3] TensorFlow, https://www.tensorflow.org/

[4] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, "Generative Adversarial Networks", arXiv:1406.2661, 2014.