

## Optical Performance Monitoring Via Histogram: A Data-Driven Approach

Yonggang Wen, Kevin Wilson

TR2009-054 September 23, 2009

### Abstract

We apply three alternative statistical learning methods to estimate optical transmission impairments (e.g., noises, chromatic dispersion) from synchronous histograms. Linear regression yields good accuracy. A more sophisticated locally weighted regression technique performs better.

*OptoElectronics and Communications Conference (OECC) 2009*

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.



# Optical Performance Monitoring via Histogram: A Data-Driven Approach

Yonggang Wen<sup>1</sup> and Kevin W. Wilson<sup>2</sup>

<sup>1</sup>Cisco Systems, San Jose, CA, USA

<sup>2</sup>Mitsubishi Electric Research Lab, Cambridge, MA, USA

Email: [ewygy@alum.mit.edu](mailto:ewygy@alum.mit.edu), [wilson@merl.com](mailto:wilson@merl.com)

(Y. Wen performed this work while at Mitsubishi Electric Research Lab.)

## Abstract

We apply three alternative statistical learning methods to estimate optical transmission impairments (e.g., noises, chromatic dispersion) from synchronous histograms. Linear regression yields good accuracy. A more sophisticated locally weighted regression technique performs better.

## Introduction

Optical performance monitoring (OPM) [1-2] will be critical for reconfigurable all-optical networks in the future. Recent research has explored performance monitoring techniques that would reside on the optical layer of these future all-optical networks [1-3]. In this paper, we focus our OPM work on quantitative assessment of two optical signal impairments, ASE noise level and fiber chromatic dispersion (CD).

In general, existing research in OPM falls into two categories [4]: i) model-based approaches and ii) data-driven approaches. The model-based approaches perform diagnosis by comparing actual observations with forecasts from an accurate model of the system. In contrast, the data-driven approaches view the system as a black box that generates characteristic patterns in its output data when a particular failure occurs. They “learn” the relationship between input events and output diagnosis in different ways, for example by using statistical methods. Kilper et al. has established the feasibility of applying data-driven approaches for OPM via data stream intensity autocorrelation in [3]. However, significant effort is needed to develop efficient statistical learning algorithms for OPM. In this paper, we exploit various data-driven methods to exploit synchronous histograms to estimate parameters of optical signal impairment (e.g., ASE noise levels and fiber chromatic dispersion coefficients). Our simulation and numerical analysis reveal that data-driven methods can estimate parameters accurately, while avoiding the complexity of developing a model for the optical transmission system.

## Data-driven Framework and Data Preprocessing

In this section, we present our data-driven framework for OPM, with an objective of accurate assessment of the signal impairments in an individual wavelength channel.

We have built a simulation testbed via the OptSim simulation system from RSoft. With it, we generate various sets of raw data to demonstrate our idea of applying data-driven approaches for investigating how well we can estimate noise levels and chromatic dispersion based solely on observations of the synchronous histogram.

Our simulation setup, as shown in Fig. 1, consists of three consecutive sections: a transmitter section, an optical link section and a receiver section. In the transmitter section, optical signals come from two sources: i) an optical source (center frequency is 193.1 THz) modulated with a 10Gbps return-to-zero (RZ) signal (for data transmission) and ii) an amplified spontaneous emission (ASE) source (for noise inclusion). The spectrum intensity of the ASE source is set to 5dBm/THz in our

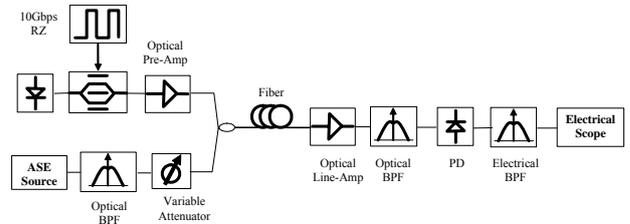


Fig. 1 Simulation setup: transmitter section, fiber link section and receiver section.

	Linear Reg	kNN	LWR
Noise Attn (dB)	1.06	0.85	0.44
CD (ps/nm-km)	0.49	0.62	0.23

Table 1: Root-mean-squared error (RMSE) summarizing the results of Figures 3. Results for kNN are for k=3 nearest neighbors.

simulation and we vary the attenuation coefficient (i.e.,  $\alpha$ ) of the variable attenuator in the range of 10 dB to 20 dB to simulate different noise levels. The optical link section consists of a pool of SMF fiber (50km) and an in-line optical amplifier (its output power is set as 6dBm). We sweep the chromatic dispersion coefficient (i.e.,  $D$ ) of the SMF fiber in the range of 5 to 20 ps/nm-km to simulate different levels of chromatic dispersion, while turning off any other non-linearity in the fiber simulation. At the receiver section, the optical signal is filtered at its specific wavelength and then is detected with a photodetector, whose output is fed into an electrical scope to acquire synchronous histograms. The data we acquire in this simulation setup is a set of synchronous histograms under different ASE noise attenuation levels and CD levels.

The dimensionality of histogram data, i.e. the number of bins, is normally very high; in practical systems, it could reach a few thousand. Therefore, the first task is to reduce the dimensionality of histogram data while extracting as much information as possible.

In optical networks, histograms are usually approximated with mixture-of-Gaussians models. As an example, for the OOK modulated transmission scheme, the histogram can be modeled with a mixture-of-Gaussians model with two components (i.e., one center for bit ZERO and the other center for bit ONE). Fig. 2 illustrates a few simulated histograms. Under this mixture-of-Gaussians model, the crucial parameters would be a quadruple  $(\mu_0, \mu_1, \sigma_0, \sigma_1)$ , where  $\mu_i$  correspond to the means and  $\sigma_i$  correspond to the standard deviations. To find the components of this mixture-of-Gaussians model, we use the expectation maximization (EM) algorithm [5], which converges to a local maximum of the likelihood function.

In addition, we obtain data from a transmission system with no noise and no chromatic dispersion and estimate its center and

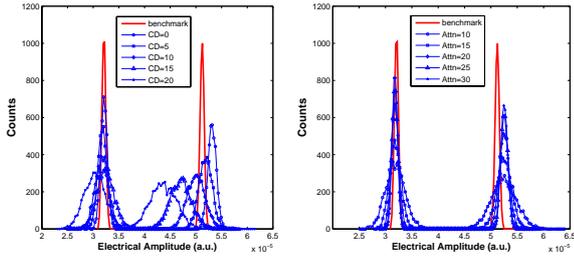


Fig. 2 Histogram comparison under different system parameters: (a) different fiber CD coefficients for a given noise attenuation and (b) different noise attenuation values for a given fiber CD coefficient.

standard deviation (i.e.,  $\mu_0^*$ ,  $\sigma_0^*$ ,  $\mu_1^*$  and  $\sigma_1^*$ ) as a benchmark. We seek to suppress the effect of specific system configuration by normalizing our test cases over the benchmark case (i.e.,  $\bar{\mu}_1 = (\mu_1 - \mu_1^*)/\mu_1^*$ ,  $\bar{\sigma}_1 = \sigma_1/\sigma_1^*$ ,  $\bar{\mu}_0 = (\mu_0 - \mu_0^*)/\mu_0^*$  and  $\bar{\sigma}_0 = \sigma_0/\sigma_0^*$ ). The normalized quadruple  $(\bar{\mu}_1, \bar{\mu}_0, \bar{\sigma}_1, \bar{\sigma}_0)$  from the data pre-processing procedure will be used as the input to the feature extraction, as described in the next section. In practical systems, the benchmark data could be obtained during the calibration phase of the network design or from a simulation testbed of the same transmission system.

#### Data-driven Performance Monitoring Algorithms

In this section, we apply three statistical learning algorithms for optical performance monitoring: linear regression (LR), nearest neighbors (NN), and locally weighted regression (LWR). Specifically, we seek to estimate both the noise attenuation level and the fiber CD level, based on the 4-D parameter vector for our mixture-of-Gaussians model. Table 1 summarizes our results in terms of root-mean-squared error (RMSE), and Fig. 3 shows the error of each technique as a function of the true noise attenuation and CD for the task of noise attenuation level estimation. In this study, we focus on the following parameter range: 5 to 20 ps/ns-km for the fiber CD level and 10 to 20 dB for the noise attenuation level. For results for all estimation techniques, we use 8-fold cross-validation.

Linear least squares regression is a standard technique that finds the best estimate of the impairment levels (i.e., the noise attenuation and the fiber CD) as a linear function of the 4-D feature (i.e.,  $(\bar{\mu}_1, \bar{\mu}_0, \bar{\sigma}_1, \bar{\sigma}_0)$ ). The  $k$ -nearest-neighbors (kNN) method [6] estimates the impairment levels from the 4-D feature (i.e.,  $(\bar{\mu}_1, \bar{\mu}_0, \bar{\sigma}_1, \bar{\sigma}_0)$ ) as the average of the  $k$  nearest neighbors to the query point in the training dataset. We tried a range of possible values for  $k$  and found that  $k=3$  nearest neighbors gave the best results (lowest error), and therefore we used  $k=3$  for our results in Fig. 3 and Table 1. To equalize the influence of each of the 4 dimensions of our input features, we scale each dimension such that it has unit variance before computing the nearest neighbors. Without doing this, a subset of the dimensions will dominate the distance calculation, and performance will be much poorer.

Finally, we apply locally weighted regression to estimate the noise attenuation and fiber CD parameters. Roughly speaking, the locally weighted regression technique is a combination of linear regression and kNN. To estimate the output value, linear regression is applied to a weighted subset of the training dataset points that are closest to the query point. We use the Locfit software package [7] to do locally linear regression. As we did for kNN, we scale each dimension to have unit variance before applying locally weighted regression.

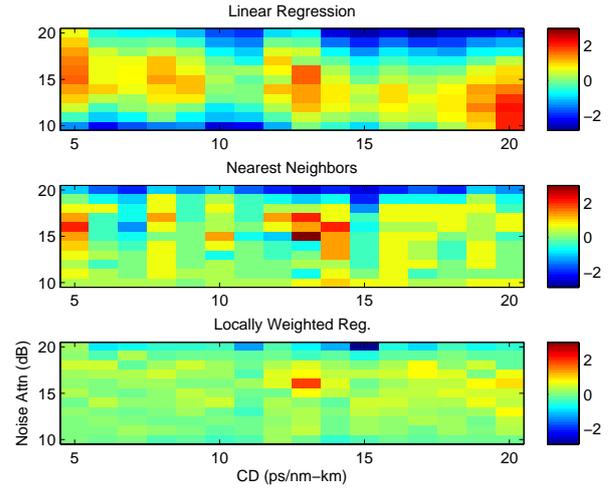


Fig. 3 Noise attenuation estimation error surface for linear regression, nearest neighbors ( $k=3$ ), and locally weighted regression as a function of true noise attenuation and chromatic dispersion. Color scale is in dB. (Error surface for chromatic dispersion is qualitatively similar but not shown due to space constraints.)

In Fig. 3 and Table 1, we compare the performance of the three methods. Note that locally weighted regression outperforms the other two techniques for both noise attenuation and fiber CD estimation. This is not surprising because locally weighted regression combines many of the advantages of both kNN and linear regression and because we have a reasonable amount of data and are working in a relatively low-dimensional (4-D) feature space.

#### Conclusion

We demonstrated the feasibility of simultaneously estimating the noise attenuation level and fiber chromatic dispersion level by extracting appropriate features from the synchronous histogram data and then applying statistical learning algorithms. Our numerical investigation indicates that the estimation is quite accurate, especially when using locally weighted regression.

#### References

- [1] D. C. Kilper, R. Bach, D. J. Blumenthal, D. Einstein, T. Landolsi, L. Ostar, M. Preiss and A. E. Willner, "Optical Performance Monitoring," IEEE Journal of Lightpath Technology, Vol. 22, NO. 1, January 2004, pp. 294-304.
- [2] L. K. Chen, M. H. Cheung and C. K. Chan, "From Optical Performance Monitoring to Optical Network Management: Research Progress and Challenges," APOC 2007.
- [3] M. Dinu, D. C. Kilper and H. R. Stuart, "Optical Performance Monitoring Using Data Stream Intensity Autocorrelation," IEEE Journal of Lightwave Technology, Vol. 24, No. 3, March 2006, pp. 1194-1202.
- [4] C. D. Cocaniala and V. Palade, "Computational Intelligence Methodologies in Fault Diagnosis: Review and State of the Art," Chapter 1 in Computational Intelligence in Fault Diagnosis, Springer London, 2006.
- [5] J. Bilmes, "A Gentle Tutorial on the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. Technical Report, University of Berkeley, ICSI-TR-97-021, 1997.
- [6] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer-Verlag, 2001.
- [7] C. Loader, "Locfit: An Introduction," Statistical Computing and Graphics Newsletter, April, 1997.