Bayesian Machine Learning and variational inference for on-site sensor calibration in Smart Bay Santa Teresa seawater monitoring

Bayesian Machine Learning and variational inference for on-site sensor calibration in Smart Bay Santa Teresa seawater monitoring

Q. Ma[1,2], G. Durin[2], F. Pennecchi[2],
C. Lombardi[3], C. Petrioli[4, 5]

[1]Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
[2]Istituto Nazionale di Ricerca Metrologica, strada delle Cacce 91, 10135 Torino, Italy
[3]Marine Environment Research Centre of S. Teresa, ENEA (Italian National Agency for New Technologies, Energy and Sustainable Economic Development), 19032 Lerici,Italy
[4]W•SENSE S.r.l., Roma, Italy
[5]University of Rome La Sapienza, Roma, Italy

E-mail (corresponding author): QitianMa@outlook.com

Keywords: Bayesian Neural Network, Sensor Calibration, Uncertainty Quantification, Seawater Monitoring

Smart Bay Santa Teresa [1] is a cooperation platform where research, technology, sustainable tourism and shellfish farming work together to make the bay an ecosystem model where Nature, man and sustainable technology interact to contrast the effects of climate change. Among various Smart Bay objectives, the platform aims to develop projects for sustainability and regeneration of Natural resources, promoting the collaborations with local entities such as small and medium-sized enterprises, administrations, up to the national and international level.

A companion project, the Green Star project, in synergy with W•SENSE [2] and Cooperativa Mitilicoltori Associati [3], aims to develop an Internet of Underwater Things (IoUT) network, built on sensors and nodes for big data production and shell farming environmental monitoring. In particular, within the Santa Teresa bay, W•SENSE measurement “nodes” have been deployed at various depths (1-5 m), consisting of sensors capable of measuring some of the fundamental oceanographic quantities (i.e. temperature, pressure, oxygen concentration, conductivity, seawater current). These measurements, carried out at a sampling rate of one data every 30 min, are transmitted by means of an acoustic modem to a gateway and then displayed and analyzed in real time. Periodically, measurements are carried out in parallel by ENEA with oceanographic instruments (e.g. a CTD probe) to verify the accuracy of the submerged measurement system.

The goal of the present activity is to calibrate on-site a W•SENSE sensor with respect to a reference CTD probe which is periodically calibrated at ENEA laboratories. A special focus is given to the measurements of oxygen concentration, which is an essential parameter for future monitoring of mussel farming production and water health state. The aim is to detect the dependence of oxygen concentration at a specific node in the grid, and correlate it with the corresponding oxygen values of the CTD and other environmental data, such as water pressure and temperature. Since the sensor is subject to the influence of environmental factors and degradation over time and the interaction among these factors is unknown, we use a Machine Learning (ML) approach able to model the W•SENSE sensor oxygen output. In addition to a point-wise calibration, uncertainty quantifications are also performed to ensure the reliability of the ML predictions [4]. Both the aleatoric and the epistemic uncertainties are evaluated, the former being caused by the noise in data, the latter (also called model uncertainty) being relevant to the variability of the model during the training phase.

Among the possible ways to address the two uncertainty contributions, we consider the Bayesian Machine Learning, in particular the Bayesian Neural Networks implementing variational inference [5]. The variational method is an analytical approximation technique to infer the posterior distribution of the model parameters by minimizing the Kullback-Leibler (KL) divergence between the true and the approximated posterior distributions. In addition, fitting a deterministic network in advance will be explored as a way to help in the decision of the prior distribution and the structure of the neural network(empiric Bayesian initialization [6]).

To evaluate the reliability of the results, we will consider different relevant metrics, such as the Mean Absolute Relative Percent Difference, the Median Absolute Error, etc..., as well as the amount of uncertainty in the ML predictions. The final aim is to develop a metrologically-sounded framework for on-site sensor calibrations able to generalize to other parameters measured in the Green Star project, as well as, possibly, to other similar sensor networks.

References

[1] https://smartbaysteresa.com/
[2] https://wsense.it
[3] http://www.mitilicoltori.it/
[4] M. Abdar et al., A review of uncertainty quantification in deep learning: Techniques, applications and challenges, Information Fusion 76:243–297 (2021)
[5] D. M. Blei et al., Variational inference: A review for statisticians, Journal of the American Statistical Association 112(518):859–877 (2017)
[6] R. Krishnan et al., Specifying Weight Priors in Bayesian Deep Neural Networks with Empirical Bayes, In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34:4477–4484 (2020)