Assessment of adaptive optics-corrected optical links statistics from integrated turbulence parameters through a Gaussian process metamodel

Emile Klotz, Sidonie Lefebvre, Nicolas Vedrenne, Christian Musso, Sylvain Poulenard, Thierry Fusco

▶ **To cite this version:**

Emile Klotz, Sidonie Lefebvre, Nicolas Vedrenne, Christian Musso, Sylvain Poulenard, et al.. Assessment of adaptive optics-corrected optical links statistics from integrated turbulence parameters through a Gaussian process metamodel. International Journal of Satellite Communications and Networking, 2024, 42 (1), pp.38-56. 10.1002/sat.1497 . hal-03805612v2

HAL Id: hal-03805612
https://hal.science/hal-03805612v2

Submitted on 21 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Check for updates

Received: 17 February 2023 | Revised: 2 August 2023 | Accepted: 16 August 2023

DOI: 10.1002/sat.1497

ORIGINAL PAPER

WILEY

Assessment of adaptive optics-corrected optical links statistics from integrated turbulence parameters through a Gaussian process metamodel

Emile Klotz[1] | Sidonie Lefebvre[1] | Nicolas Vedrenne[1] | Christian Musso[1] | Sylvain Poulenard[2] | Thierry Fusco[1,3]

[1]DOTA, ONERA, Paris-Saclay University, Paris, France

[2]Airbus Defence and Space, Toulouse, France

[3]Laboratoire d'Astrophysique de Marseille, Aix Marseille Univ, CNRS, CNES, Marseille, France

Correspondence
Emile Klotz, DOTA, ONERA, Paris-Saclay University Fr-91120, France.
Email: emile.klotz@onera.fr

Funding information
This research was done in the framework of Emile Klotz's PhD thesis partially founded by Airbus Defence and Space.

Summary

With the development of free space optical links (FSO) for space-ground communications comes the need to mitigate the effects of the atmospheric turbulence to guarantee a lossless connection. By having a network of addressable ground stations, we want to guarantee to always target a point where the link is available. Assuming atmospheric transmission is managed thanks to site diversity, we focus only on the influence of atmospheric turbulence on the signal injected into a single mode fiber on the downlink. The use of adaptive optics (AO) is assumed to avoid turbulence-induced signal disruptions and enable a sufficiently high level of received signal for data transmission. Up to now, AO performance adequate assessment required the knowledge of high-resolution C_n^2 and wind profiles. With the advent of integrated atmospheric parameters measurement instruments, we investigate here the possibility to estimate AO-corrected performance from a limited number of integrated parameters. In this paper, we propose to use a Gaussian process metamodel to assess the statistics of the received optical power after an AO correction. Taking as input only four integrated parameters of the turbulence profile and associated wind profile, which can be measured with simple instruments, the estimation error on the value of the 1% quantile of the received optical power is inferior to 0.7dB. We also demonstrate the possibility to estimate the half correlation time of the received optical power using the same integrated parameters.

KEYWORDS
adaptive optics, FSO, machine learning

1 | INTRODUCTION

To match the growing need for data transmission between the ground and space that is driven by the development of space data highways[1] and upcoming mega communication constellations,[2] optical links become an increasingly credible alternative to radiofrequency links. Offering a favorable size, weight and power, and frequency allocation free and intrinsically secured very high data rate transmission, the implementation of optical communication between space and the ground is the subject of sustained developments. Many projects are ongoing to establish such a link

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. *International Journal of Satellite Communications and Networking* published by John Wiley & Sons Ltd.

between a GEO satellite and a fixed Earth ground station for missions such as Internet delivery and/or data repatriation from point to point on earth. However, the development of these atmospheric optical links remains conditioned by their availability, which is highly dependent on the atmospheric channel: absorption, scattering, and turbulence. Cloud masking issues can be managed thanks to site diversity,[3] and we focus here on atmospheric turbulence influence.

With the wish for very high bandwidth comes the necessity to inject the signal into a single mode optical fiber (SMF) to be amplified and/or to enable coherent detection. Deep fluctuations in the injected signal happen due to atmospheric turbulence that causes amplitude variations, that is, scintillation and wavefront distortions. A significant telescope aperture enables averaging of the scintillation effects as the size of the speckles becomes small in comparison with the receiving aperture. Phase-related fadings in the received signal can be reduced with the use of adaptive optics (AO) systems to correct the distorted wavefront and maximize the coupling efficiency in the SMF.[4,5] AO and pupil averaging, because of technological and cost limitation, does not perfectly correct for turbulent channel errors. To maximize the retrieved information, digital mitigation techniques can be used such as forward error correction on an interleaved signal[6] in addition to the physical mitigation technics. Forward error coding will operate efficiently if the error probability (hence fading probability) remains reasonable on the interleaved signal. The size of the interleaving window depends on temporal characteristics of the received optical power (ROP) such as the typical coherence time.[7] Those techniques will have a big impact on the bandwidth and on the latency of the whole system and thus must be adequately scaled according to the ROP. Knowing what would be the ROP distribution and ROP coherence time on any site and at any moment would give us precious information on how to adapt the system in real time. However, this statistic highly depends on turbulence conditions knowledge along the line of sight and on the AO system. Considering the significant variability of turbulence conditions with the location of ground stations, several initiatives have recently been taken to gather local measurements of the most relevant atmospheric parameters with respect to AO-corrected optical links. Some of these initiatives deploy high vertical spatial resolution measurements,[8] whereas others focus on integrated turbulence parameters.[9]

Knowledge of the turbulence and wind profiles at kilometer vertical resolution along the line of sight guarantees to precisely assess fading statistics,[10] hence an accurate optical link performance evaluation. High-resolution atmospheric turbulence characterization can be performed thanks to rather complex instrumentation such as Moon Limb Profiler by nighttime[11] or Sun Limb Profiler.[11] Such instruments are being deployed to demonstrate high-resolution turbulence characterization capacity on several sites over Europe.[12] They will provide high-resolution C_n^2 profile characterization along the line of sight between the ground and the direction of the target used, which might be different from the optical links direction, the impact of this difference of line of sight on the optical links availability being hardly documented. Initiatives to systematically document integrated turbulence parameters sometimes with really simple instrumentation are also emerging.[13] Considering the importance of integrated turbulence parameters in the assessment of AO-corrected error budgets, indications exist in favor of a link performance which would depend only on a few integrated parameters, but such a relation between corrected optical link performance and integrated turbulence parameters has never been clearly established so far. The exact expression of the correction residuals involves a complex combination of moments of the C_n^2 profile and wind, whose weightings depend on the number of corrected modes and the tractability of analytical expressions raises real challenges for a clear-cut demonstration that integrated parameters are sufficient to characterize optical link availability. Machine learning methods associated to physical performance models might provide crucial indications to answer this question.

It is the major prospect of this paper: investigating the possibility to assess AO-corrected optical link availability from integrated turbulence parameters by machine learning and to identify the compulsory parameters to be monitored, thanks to a sensitivity analysis.

Over the past 10 years, some studies have taken advantage of machine learning for atmospheric turbulence estimation or temporal prediction. Most focus on assessing C_n^2 near the land surface, such as Wang and Basu[14,15] which propose to use a multilayer perceptron (MLP) trained, respectively, on seven measured meteorological input variables: wind speed, temperature and temperature gradient, soil temperature, relative humidity, net radiation and sol water content, or only five input variables: wind speed, relative humidity, pressure, wind shear, and potential temperature gradient. In Su et al,[16] only four meteorological variables are used: temperature of the surface, temperature, wind speed, and relative humidity measured at 0.5 and 2 m. Prediction results are overall accurate but associated with a particular scenario. In addition to these MLP metamodels, Jellen et al[17] compared three other metamodels: polynomial regression, random forest, and boosted regression trees with six input variables: air temperature, air-water temperature difference, pressure, relative humidity, wind speed, and solar radiation. Best results were obtained with random forest, but the predictions were not always accurate. Some deep neural networks have also been used more recently: in Lamprecht et al,[18] a ResNet residual network is proposed to retrieve the refractive index structure parameter from the height above sea level and the corresponding wind speed, instead of relying on analytical formulae. The performances are promising, but it would require to collect training data from many different places on earth in order to deliver accurate results. A recent PhD work aimed at forecasting future daytime C_n^2 conditions from prior meteorological data: wind speed, pressure, temperature, relative humidity and solar irradiance, and C_n^2 measurements.[19] Neural networks (MLP and recurrent neural network) are used to create a low-altitude model capable of forecasting C_n^2 up to 4 h later using 16 h of prior measurements. The forecasting quality is not always sufficient, best in the middle of the day, moderate in the morning, and generally worst in the evening.

Finally, some recent approaches[20] use deep neural network to infer the atmospheric turbulence refractive index structure parameter C_n^2 from short exposure images of turbulence-induced laser beam intensity scintillations.

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Other studies focus instead on integrated turbulence parameters temporal prediction. Among these, we can cite Milli etal.,[21] where turbulence nowcasting, that is, the ability to forecast the turbulence conditions over the next 2 h, is investigated at Paranal Observatory. MLP performed best among the three metamodels tested: random forest, MLP, and long short-term memory (LSTM) deep network trained on 1 or 2 h history of meteorological and integrated turbulence parameters such as seeing, coherence time, temperature, pressure, wind speed, and direction. In Giordano et al,[22] a random forest metamodel is trained to predict the seeing over the next 2 h on a large atmospheric database measured by the Calern Atmospheric Turbulence Station, including ground meteorological conditions, vertical profiles of the C_n^2, and integrated parameters characterizing the optical turbulence: seeing, isoplanatic angle, and coherence time.

As far as free space optical links (FSO) are concerned, machine and deep learning methodology mostly focus on compensating the effects of atmospheric turbulence on the performance of the whole single input single output (or SISO) and multiple input multiple output (or MIMO) FSO system[23] or on predicting parameters of the FSO channel,[24] such as optical signal-to-noise ratio. Closer to the methodology we propose in this paper, two publications aimed at predicting the RSSI (received signal strength indicator) of the FSO. In Tóth et al,[25] pressure, air temperature, particle concentration, visibility, relative humidity, and wind speed at different past time horizons are used as input variables for the metamodels. Best results were obtained with random forest and enabled to retrieve some atmosphere behavioral patterns influencing RSSI. Lionis et al[26] compared different metamodels: k-nearest neighbors, tree-based methods-decision trees, random forest, gradient boosting, and MLP trained on seven local atmospheric parameters: wind speed, pressure, temperature, humidity, dew point, solar flux, and air-sea temperature difference. Best determination coefficient R_2 is 0.949 and is obtained with the MLP metamodel.

Here, we focus on the injected power statistics assessment, and we consider input variables associated to the turbulence profile and to the wind profile in a perspective to identify a minimum of compulsory parameters. Using machine learning on a database of turbulence profiles over Tenerife and a physical AO modeling tool, we propose a metamodel to estimate the probability density function (PDF) of the injected power into the single mode fiber of the receiver of the optical ground station. We further develop this metamodel to be able to assess the autocorrelation function of the injected power. For our method to be suitable for a massive deployment of ground stations around the world, we wish to use exclusively data provided by simple and easily deployable instruments.

Section 2 describes our atmospheric channel model and the profiles database, Section 3 presents the methodology, including metamodel construction and sensitivity analysis, and Section 4 discusses the numerical results.

2 | ATMOSPHERIC CHANNEL MODELING

The metric of free space optical communication availability in the case of a coded channel is given by Shannon's noisy-channel coding theorem[27] that states that for successful decoding with arbitrarily small error probability, the capacity of a communication channel must be greater than the rate of the code used (the proportion in bits of the datastream that is useful, i.e., nonredundant). Unfortunately, due to slow-fading or other random factors, the channel capacity can fall below the code rate, hence compromising error-free decoding. In other words, the probability of interruption of the turbulent channel is $P_{outage}(R_0) = P(C(SNR) < R_0)$, where C is the channel capacity defined by Shannon[27] as the maximal data rate that can be achieved with the given channel, SNR is the electrical signal-to-noise ratio, and R_0 is the coding rate of the forward error correction code.

The capacity and, on a wider scale, the SNR are strongly dependent on the reception system. The challenge of accounting for all the intricate phenomena arising from different types of receiving chains is beyond the scope of this article. We present here a proof-of-concept study in a straightforward scenario, and we focus our attention on the received power into a SMF (the ROP) statistics. Looking at the ROP statistics will allow us to take into account the interleaving and the coding while avoiding strong hypothesis on the noise statistic, which allows a more exhaustive approach. In the very simplistic case of a noise that would be independent of the signal, we can show that there exists a bijection between the channel capacity fluctuations and the fluctuations of the ROP.[28]

In order to train our metamodel to estimate the ROP statistics, we first need access to a database of ROP and associated turbulence measurements. All those data are obtained thanks to simulations that we want to be as representative as possible of the most significant phenomena that affect AO performance in the limit of tractability of the model, in the prospect to build a methodology and identify the most adapted ML tools to extract relevant informations from the data. A necessary adaptation of the inputs used in the metamodel will be needed when switching to experimental data. In this section, we describe the AO correction modeling, the database, and the system hypotheses.

2.1 | Modeling of the ROP after AO correction

In this work, we study separately the impact of the scintillation on the ROP and the impact of the distorted phase. We call ρ_ϕ the coupling efficiency neglecting the impact of scintillation and ρ_I the term of scintillation. Correlated time series of ρ_ϕ and ρ_I are obtained by modeling the fluctuation of the coupled flux in a SMF with a pseudo-analytic AO modeling tool called SAOST (simplified adaptive optics simulation tool).[10,29,30] This

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

model neglects the influence of noncommon path aberrations between the wavefront sensor and the injection path and assumes a perfect wavefront sensor (not sensitive to scintillation), a Zernike description of the correction phase, and an infinitely fast deformable mirror (the delay in the loop mostly comes from integration time of the wavefront sensor and calculation of the control voltages sent to the mirror). Comparison between SAOST and end-to-end models can be found in Canuet.[31] Experimental validation has been conducted in relevant condition for a GEO feeder link and can be found in Bonnefois et al.[32]

In case of moderate turbulence strength, terms of interaction between the scintillation, mostly caused by distant turbulence, and the phase effects, related to close to ground phenomena, can be neglected, and thus, SAOST works under the assumption of independence between ρ_ϕ and ρ_I. Let $f_{SMF} = \rho_\phi * \rho_I$ be the coupling efficiency in a SMF after propagation through atmosphere and AO correction. Analytical expressions of f_{SMF} and its distribution under the hypothesis of independence are described in more details in Canuet et al.[10] Scintillation influence is simulated assuming the small perturbation approximation in the Rytov regime. In practice, this constrains the validity of the approach to limited Rytov variance (typically when the point source log-amplitude variance $\sigma_\chi^2 < 0.3$) which corresponds to the experimental limit of validity of the small perturbation approximation for horizontal propagation as first identified in Gracheva and Gurvich.[33] This strong limitation in the case of horizontal propagation is, to our understanding, not restrictive when dealing with vertical propagation where σ_χ^2 takes low values. In the database presented in the following part containing 37,059 profiles, less than 0.05% have a $\sigma_\chi^2 \geq 0.3$.

We adopt in this paper the same approach as in SAOST, that is, to consider independently the effects due to the phase error and those due to the scintillation; the benefit of this dissociation is to be able to interpret our results more easily.

2.1.1 | Residual phase error model

The residual phase error is computed in SAOST using a Monte Carlo approach. Following the algorithm described in Roddier,[34] random occurrences of Zernike coefficients are sampled to describe the corrected phase. The temporal correlation of the ROP time series is obtained by filtering the raw Zernike coefficients by a temporal power spectral model[35] in the Fourier domain. Some hypothesis have to be made on the turbulence condition of which we can cite von Kármán statistics for the index of refraction fluctuations spectrum assuming a fixed outer scale. More details and comparisons with an end-to-end simulation can be found in Canuet.[31]

2.1.2 | Scintillation impact

The temporal impact of scintillation on the coupling efficiency can be approximated by the product[10]

$$\rho_I = \exp\left(-\sigma_\chi^2\right) \exp(2\chi_{AP}(t)) \tag{1}$$

in the weak fluctuation regime of the scintillation index $\chi(r,t)$ where σ_χ^2 is the variance of the punctual log-amplitude and $\chi_{AP}(t)$ is the log-amplitude averaged on the receiver aperture.

Under weak irradiance fluctuation hypothesis, we have $\chi_{AP}(t) \sim \mathcal{N}\left(-\sigma_{\chi_{AP}}^2, \sigma_{\chi_{AP}}^2\right)$. We can then write the normalized irradiance distribution, $\exp(2\chi_{AP}(t))$, as the following log-normal distribution[36,37]:

$$f_{\rho_I}(x) = \frac{1}{2x\sigma_{\chi_{AP}}\sqrt{2\pi}} \exp\left[-\frac{\left[\ln(x) + 2\sigma_{\chi_{AP}}^2\right]^2}{2\left(2\sigma_{\chi_{AP}}\right)^2}\right], x > 0 \tag{2}$$

2.2 | System hypothesis and training database

2.2.1 | AO and simulation parameters

In the following, all the generations of power attenuation's time series will be done with the same parameters of AO. We assume an AO system that corrects the first 10 radial orders with a frequency of 2kHz. The simulation is done with a time sampling of 4kHz for a duration of 10 s which

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

gives us 40.000 points per time series. Our telescope is taken with a pupil of 60cm. These choices are being made as they assume a relatively simple and cost effective hardware for this type of application.

2.2.2 | Atmospheric conditions database

As we want to describe the ROP statistics using a set of integrated parameters from the C_n^2 and wind profile, the best approach is to build a large database of such profiles.

It should be noted that knowledge of the C_n^2 and wind profiles at kilometer vertical resolutions guarantees an accurate optical link performance evaluation using simulation tools such as end-to-end modeling of the atmospheric propagation and the correction by AO. Unfortunately, it is challenging to determine the profile of C_n^2 with altitude even with complex and expensive instruments such as SLODAR (slope detection and ranging), SCIDAR (scintillation detection and ranging) and SODAR (sonic detection and ranging).[38–40] In addition to this, these profiles present the risk to be affected by measurement noise (hence, influencing the metamodel). To our knowledge, there is no experimental database available that would be representative of the various atmospheric conditions, which leads us to use a database from an atmospheric reanalysis model. Indeed, data provided by numerical models present the advantage to precisely control underlying hypothesis and input parameters at the expense of a more questionable relevance. As the first goal of this work being to demonstrate the possibility to rely on a metamodel for performance assessment in relevant conditions, the possibility to cover a large scope of atmospheric conditions justifies in itself to exploit data obtained from a numerical model.

There exist models based on empirical measurements, some nonparametric such as Greenwood's,[41] H-V Night,[42] and AFGL AMOS[43] and some parametric such as the famous Hufnagel–Valley model,[44] its enhanced version the Hufnagel–Andrews–Phillips[45] model, or the Sadot–Kopeika[46] model. All those models are suitable to describe the average value of the turbulence over a given site and are useful for site selection, but they will not give an instantaneous and accurate turbulence profile description. Overall, these models have been developed for specific sites at given time of the year, and it is not clear to what extent they would be suitable to other sites or meteorological conditions.

As our work aims at characterizing the statistics of the instantaneous ROP and, especially, the distribution's tail to describe the probability of interruption, we need to work with a theoretical model that would describe any small variation in the turbulence induced by different meteorological conditions and would work for any location and any hour of the day.

Such a description is achieved using Gladstone's formula and Tatarskii's[36] theory with a model for the outer scale such as Dewan's[47] or HMNSP99.[48] In this approach, the C_n^2 profile is calculated from precise vertical profiles of meteorological parameters (temperature, pressure, relative humidity, wind speed, and wind direction).

Our database of profiles was provided by Durham University. Wind and C_n^2 profiles were obtained through a global turbulence model capable of converting meteorological data from a general circulation model, into three-dimensional optical turbulence maps. This model based on Tatarskii's is developed in Osborn and Sarazin[49] and was confronted successfully to on site measurements in Paranal. It was improved to include a separate boundary layer and enable stronger turbulence strength near the ground to be modeled.[50]

The general circulation model used is ERA5[51] from the European Centre for Medium range Weather Forecasts (ECMWF). This model, from which the turbulence is calculated, has a spatial resolution of 0.3° along latitude and longitude and provides a forecast for every hour. We chose a grid of 11 by 11 points around Tenerife's island (with a spatial resolution of $\approx$ 30 km), which is a site of interest for a potential future ground station, and focused on the first 19 days of March 2018. It leaves us with 121 simultaneous measurements for each hour, with some missing values. On the overall 19 days considered, we thus have 37,059 profiles on 113 pressure levels each. Historical data are freely available on ECMWF website.

These 37,059 C_n^2 and wind profiles cover a large set of condition. In the following, we consider that these profiles are representative of field data and could be obtained with instruments directly measuring turbulence along the line of sight.

3 | MACHINE LEARNING METHODOLOGY

As described in Sections 2.1, we consider the residual phase error and the scintillation as two independent phenomenons and the quantity of interest, the coupled flux in a single mode fiber, as the product of both.

In Sections 2.1.2, we presented our scintillation simulation model, which is based on a parametric expression of the distribution that depends only on the two parameters σ_χ^2 and $\sigma_{\chi_{AP}}^2$ (Equation (1)). However, the latter is not easily measurable, as it would require an instrument with the same pupil size as our telescope. Our first goal is therefore to be able to assess $\sigma_{\chi_{AP}}^2$ with machine learning using a small number of easily measurable parameters. Our second goal is to describe the PDF of ρ_ϕ with the minimum amount of variables and be able to determine those variables from the same small number of measurable parameters.

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

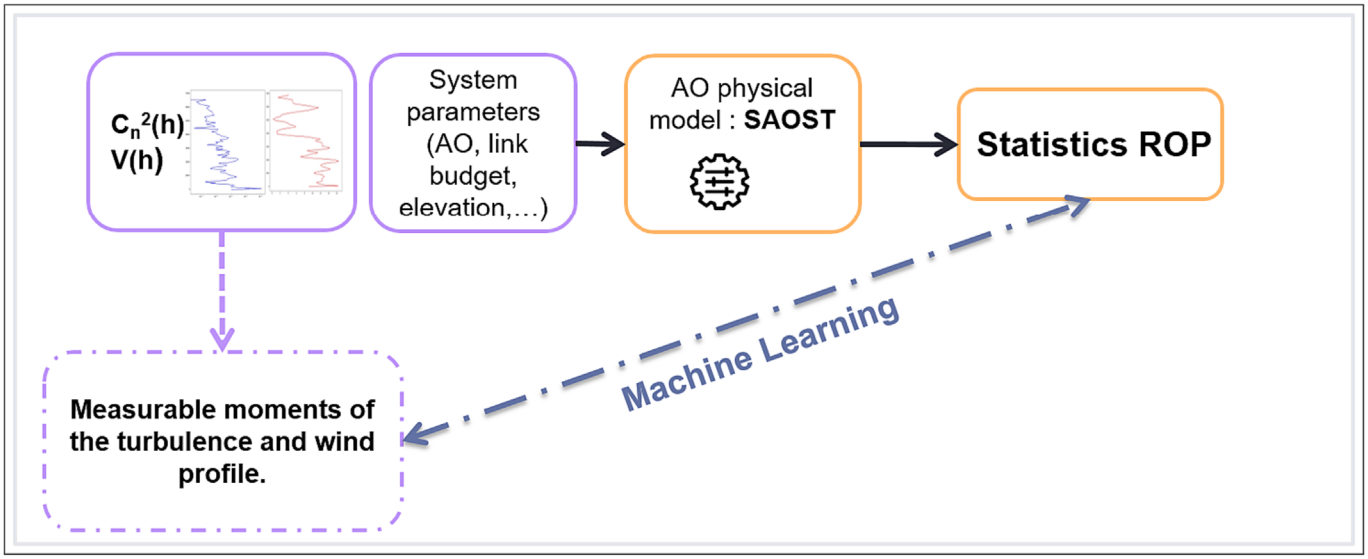

FIGURE 1 Schematic description of the method proposed in this paper.

Figure 1 sums up the methodology; the upper boxes with full lines represent the current state of modeling the statistic of ROP. It requires the use of C_n^2 and wind profiles coupled with a physical model of light propagation through the atmosphere and wavefront correction with an AO loop. What we aimed at is represented by the dashed box and arrows as we want to shortcut the heavy process of profiles measurement and simulation thanks to machine learning on a small number of integrated parameters easily measurable.

3.1 | Residual phase error distribution description

A usual and effective way to predict a distribution using machine learning is to parameterize this distribution and then use a metamodel to predict each parameter.

Here, we were led to study

$$L_\phi(t) = 10\log_{10}(\rho_\phi(t)) \tag{3}$$

the loss in power induced by the phase fluctuation in dB as it gives a bigger weight to the low values of $\rho_\phi(t)$ that are the critical values for our application.

Studying the distribution of $\rho_\phi(t)$ on our 37 k profiles, we highlighted that it has an exponential decay, which is consistent with the closed form of the distribution already proposed by Canuet.[31] We can show that the logarithm of an exponential distribution is a Gumbel distribution, and we verified experimentally that the distribution that best fit the distribution of our loss $L_\phi(t)$, in perspective with the Bayesian information criterion and sum of square error, is a Gumbel distribution of the following form:

$$\frac{1}{\beta} e^{-(z+e^{-z})} \tag{4}$$

where $z = \frac{x-\mu}{\beta}$.

This result is particularly interesting as, with the Gumbel distribution being a good enough approximation of the density probability of power attenuation, we can describe the PDF with only two parameters, μ and β, that contain all the information on the statistic of power attenuation.

The quality of the fit can be seen by looking at the relative error measured between the quantiles of the experimental distribution and those of the theoretical distribution. We show in Table 1 the statistics of relative error made on some relevant quantiles. The statistics are given for the whole database. For example, the relative error on the 0.01 quantile is, in average, of 0.57% and is below 1.92% in 99% of the cases. Such small relative errors emphasize to which extent our fit is appropriate. It is to be noticed that, the number of data of our simulations being finite, the error on the smallest quantiles can be due to a lack of data to precisely estimate the latest, as much as a nonability of the Gumbel distribution to

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Relative errors on quantiles with the Gumbel fit over the 37,059 power attenuation's series.

	Quantile				
Relative error	**0.001**	**0.003**	**0.01**	**0.03**	**0.1**
Mean	1.43%	0.91%	0.57%	0.37%	0.19%
1%	0.02%	0.01%	0.01%	0.01%	0.01%
50%	1.17%	0.75%	0.48%	0.33%	0.17%
99%	5.36%	3.28%	1.92%	1.14%	0.54%

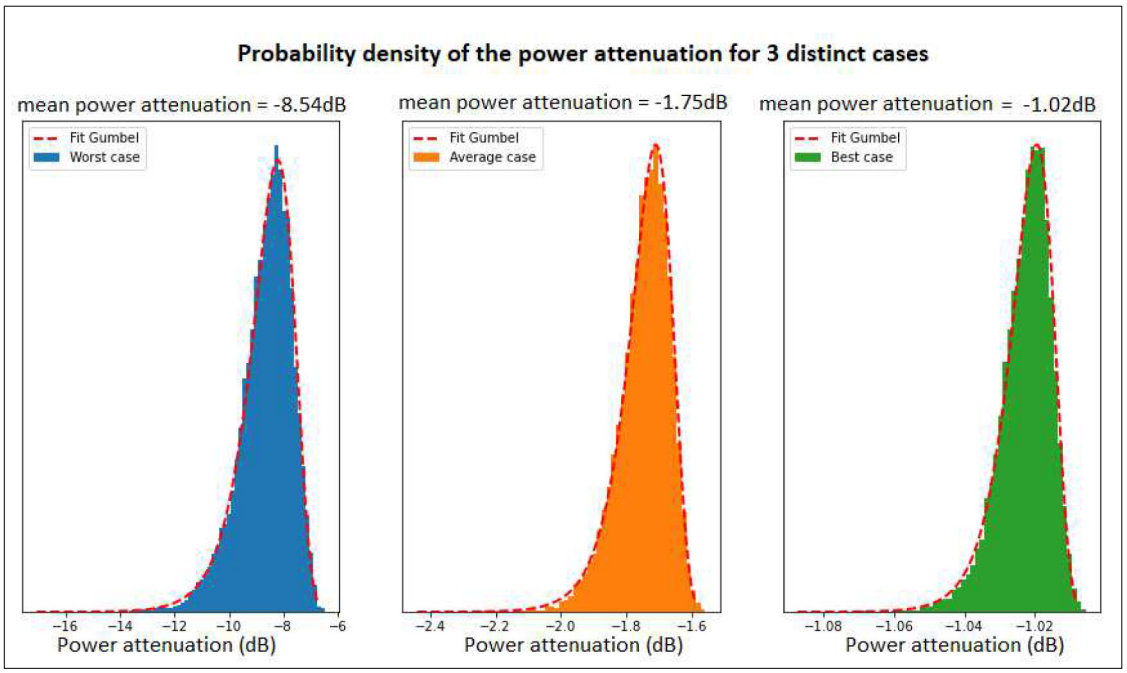

FIGURE 2 Gumbel fit on the distribution of L_ϕ in three cases: First one in blue corresponds to the minimum of coupling efficiency, second one in orange is an average coupling efficiency, and the last one in green is obtained for the profile with the highest coupling efficiency.

describe precisely the smallest quantiles. Further work in this regard was not conducted as the quality of the approximation is good enough for our application.

Figure 2 aims at visually demonstrating the goodness of the fit in the diversity of conditions encountered. We selected three sets of C_n^2 and wind profiles from the whole database, one corresponding to the average value of mean power attenuation over the 37,059 profiles (orange), one to the worst value of mean power attenuation (blue), and the last one to the best value of mean power attenuation (green).

3.2 | Gaussian process (GP) metamodel

Different machine learning techniques have been tested of which we can cite gradient boosting, MLP, support vector machine regression, and GPs.[52] Best results were obtained for a GP, and we only describe this metamodel here.

A GP writes the output of interest as the sum of a regression part, a constant term in this study, and a centered stochastic process Z:

$$\gamma(x) = \beta_0 + Z(\mathbf{x}) \tag{5}$$

The stochastic part $Z(\mathbf{x})$ is a Gaussian-centered process fully characterized by its covariance function $Cov(Z(\mathbf{x}),Z(\mathbf{u})) = \sigma^2 R(\mathbf{x},\mathbf{u})$ with σ^2 the variance of Z and R the correlation function, or kernel, that accounts for spatial correlation effects.

In this study, we focus on a stationary process Z, which means that, for new points, the prediction consists of a linear combination of the observed values, with weights that depend on the distance between the new input point and the training data. The assumption is that the closer the inputs are, the more correlated the outputs are. The kernel we chose is a Matérn 5/2 kernel, an extension of the radial basis function kernel,

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

one of the most commonly used forms of kernel. The Matérn kernel computes the similarity of two given points $\mathbf{x}$ and $\mathbf{x}'$ in dimension d as follows:

$$R(\mathbf{x}-\mathbf{x}') = 1+\sqrt{5}\sum_{j=1}^{d}\frac{\left(x_j-x_j'\right)}{\theta_j}+ \frac{5}{3}\sum_{j=1}^{d}\frac{\left(x_j-x_j'\right)^2}{\theta_j^2}\exp\left(-\sqrt{5}\sum_{j=1}^{d}\frac{\left(x_j-x_j'\right)}{\theta_j}\right) \tag{6}$$

where the θ_j are the hyperparameters and should be optimized in addition to β_0.

One of the advantages of GPs is that an estimate of the uncertainty associated with the prediction is available.

In order to evaluate the predictive ability of the metamodel, we rely on the predictivity coefficient Q_2, which stands for the percentage of the output variance explained by the metamodel. It is the same as the determination coefficient R_2 but computed on n test data Y_i instead of training ones:

$$Q_2 = 1-\frac{\sum_{i=1}^{n}\left(Y_i-\widehat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2} \tag{7}$$

where $\overline{Y}$ is the mean of the test data and $\widehat{Y}_i$ stands for the output of the metamodel for the same input values as Y_i. Q_2 is between 0 and 1 and should be close to 1 for an accurate metamodel prediction.

3.3 | Choice of relevant inputs

We have based our choice of integrated parameters on the integrated parameters regularly used to describe the error budget of an AO. The first chosen parameter is Fried's parameter, an essential parameter when one is interested in the effects of turbulence. Fried's parameter r_0 is defined as the typical diameter of a telescope whose resolution would be limited by atmospheric turbulence. In the case of a plane wave, considering a Kolmogorov spectrum, we have along the line of sight:

$$r_0 = \left[0.42\left(\frac{2\pi}{\lambda}\right)^2\int_0^{\infty}C_n^2(z)dz\right]^{-\frac{3}{5}} \tag{8}$$

where λ is the wavelength (1.55 μm). Experimentally, Fried's parameter can be estimated either from the amplitude jitter of a star at the focal plane of an imager or more robustly thanks to differential imaging such as performed with a DIMM.[53]

The second parameter, denoted $\overline{h}$,[54] is a measure of the height dispersion of atmospheric layers, homogeneous to an altitude:

$$\overline{h} = \left[\frac{\int_0^{\infty}z^{\frac{5}{3}}C_n^2(z)dz}{\int_0^{\infty}C_n^2(z)dz}\right]^{\frac{3}{5}} \tag{9}$$

It provides an assessment of the physical origin for the angular decorrelation of the phase perturbations (the influence of distant turbulence layers) while being independent from the turbulence strength. It is useful to characterize and compare different profiles. $\overline{h}$ is related to Fried's parameter and the isoplanatic path θ_0 by[55]:

$$\theta_0 = 0.314\frac{r_0}{\overline{h}} \tag{10}$$

The estimation of $\overline{h}$ can therefore be performed thanks to a measurement of θ_0, which is derived, for instance, from limited aperture averaged scintillation by nighttime[56] or thanks to a Shabar measurement by daytime.[57] The last parameter, denoted $\overline{v}$, describes an average wind speed over the turbulent layers[59] and similar to $\overline{h}$ is given as follows:

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\bar{v} = \left[\frac{\int_0^\infty v(z)^{\frac{5}{3}} C_n^2(z) dz}{\int_0^\infty C_n^2(z) dz} \right]^{\frac{3}{5}}, \tag{11}$$

where v is the modulus of the transverse wind velocity. It is related to the turbulence coherence time τ according to the following:

$$\tau = 0.314 \frac{r_0}{\bar{v}} \tag{12}$$

The turbulence coherence time can be extracted, for instance, from the temporal analysis of the jitter of a bright enough point source image. This evaluation can then be exploited to provide an estimation of $\bar{v}$.

r_0, $\bar{h}$, and $\bar{v}$ are quantities that characterize the phase of the wave and σ_χ^2 the intensity fluctuations. The link between phase and intensity perturbations involves diffraction, which is not accounted for in the calculation of the integrated parameters, hence the need to be able to measure σ_χ^2

The Rytov approximation discussed in Section 2.1 assumes that the refractive index fluctuations are small compared with the mean refractive index, allowing for linearization of the wave propagation equation. This approximation is valid for weak turbulence, where the fluctuations are small and the wavefronts do not experience significant bending. Under this regime, σ_χ^2 can be approximated in the following way:

$$\sigma_\chi^2 = 0.563 \left(\frac{2\pi}{\lambda} \right)^{\frac{7}{6}} \int_0^\infty z^{\frac{5}{6}} C_n^2(z) dz \tag{13}$$

According to,[33,58] σ_χ^2 value is precise when it is less than 0.3, which is in accordance with the studied database.

Based on this considerations, the point source log-amplitude variance σ_χ^2 is supposed to be measured from the scintillation of a bright point source with a small diameter instrument.

Thus, the PDF of the scintillation only depends on the unknown parameter $\sigma_{\chi_{AP}}^2$.

3.4 | Sensitivity analysis

In this study, we rely on Sobol's indices,[60] also known as variance-based sensitivity analysis. The variance of the output of interest Y is decomposed into fractions, which can be attributed to each of the moments we use here as input variables for our metamodel. The values of Sobol's indices enable to rank input variables according to their importance in the uncertainty of the output. The first-order Sobol index S_i characterizes the contribution of a given input X_i to the output variance, and the total Sobol index S_{T_i} measures the contribution to the output variance of the studied input X_i, including all variance caused by its interactions of any order with all other input variables. They are estimated thanks to a Monte Carlo method and can be written as follows:

$$S_i = \frac{Var(E[Y|X_i])}{Var(Y)} \tag{14}$$

and

$$S_{T_i} = 1 - \frac{Var(E[Y|X_{\sim i}])}{Var(Y)} \tag{15}$$

where E stands for the esperance of the random variables and $X_{\sim i}$ all X except X_i. First-order indices vary between 0 and 1, and the difference between 1 and their sum characterizes the global influence of interaction effects. If the total index associated to an input variable is close to zero, this input has a negligible impact on the output variability and can be set at a constant value. On the contrary, Sobol's indices close to one indicate that the input variable is influent.

Sobol' indices are very often used to determine the sensitivity of a simulation code to a specific input but works under the assumption of independence between input variables; their interpretation becomes hazardous in the case of correlated inputs. In our case, r_0, $\bar{h}$, and $\bar{v}$ are strongly correlated being all moments of the same profiles. To deal with correlated inputs, methods have been developed around Shapley values, which come from the field of cooperative games theory.[61,62] The associated Shapley indices are designed as a simple and easy way to interpret effects of the interactions and dependences contributions between the inputs involved on the total output variance.

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4 | NUMERICAL RESULTS

4.1 | Probability density assessment

4.1.1 | Metamodel construction

Our database is split in two datasets: one training set containing 15% of the profiles randomly selected and a test set containing the remaining 85%. This arbitrary choice corresponds to 5558 randomly selected profiles. Less profiles would be enough to learn the few parameters of this GP as long as they cover the variability of the encountered cases.

The complexity of GP algorithms is $O(n^3)$ due to the need to invert an $n \times n$ matrix; a 15% training set is a good trade-off between prediction accuracy and learning time with 20 min of single core time on a modern processor for the training process.

The input vector of Equation (5) is $\mathbf{x} = (r_0, \overline{h}, \overline{v})$.

Figure 3 shows the estimated value against the one obtained with our metamodel. The coefficient of determination Q^2 is superior to 0.99 in every case and shows how observed outcomes are predicted by the model. We find 0.9994 for the prediction of μ, 0.9992 for β and 0.9969 for $\sigma_{\chi_{AP}}$.

As we are able to predict μ, β, and $\sigma_{\chi_{AP}}$ from the moments, the next step is to reconstruct the probability density of the ROP using the parametric descriptions of $L_\phi(t)$ (Equation (3)) and ρ_l (Equation (1)). An example is given in Figure 4 where, for one randomly selected profile, we can see, from left to right, the statistical reconstruction of $L_\phi(t)$, ρ_l, and the ROP.

In order to characterize the relevance of our estimation on the ROP's statistic, we can compute the absolute error made on the mean and standard deviation of the reconstructed PDF of the ROP. We also look at the absolute error on the 1% quantile of the ROP (see red dashed line in Figure 4) as we want a faithful reproduction of the tail of the distribution. Statistics on the absolute error associated to our 37 k profiles can be seen in Figure 5.

On all profiles, the prediction error on the value of the 1% quantile is inferior to 0.7 dB which is compatible with current assumptions done in commonly used link budgets (margins are typically 3dB).

Once again, we have to put into perspective this value with the fact that the temporal series of ROP generated with SAOST are finite, and thus, part of the error is due to the nonperfect convergence of the random variable. The weight of this error due to convergence in the overall error has yet to be determined.

4.1.2 | Sensitivity analysis

After finding satisfactory results with our metamodel, we want to describe the impact of each input variable on each output. To do so, we are using sensitivity indices: first and total order Sobol indices as well as Shapley indices, which were presented in Sections 3.4. Sensitivity indices were estimated using "Sensitivity: Global Sensitivity Analysis of Model Outputs",[63] an open source, GPL-2 licensed, R[64] library developed for the treatment of uncertainties.

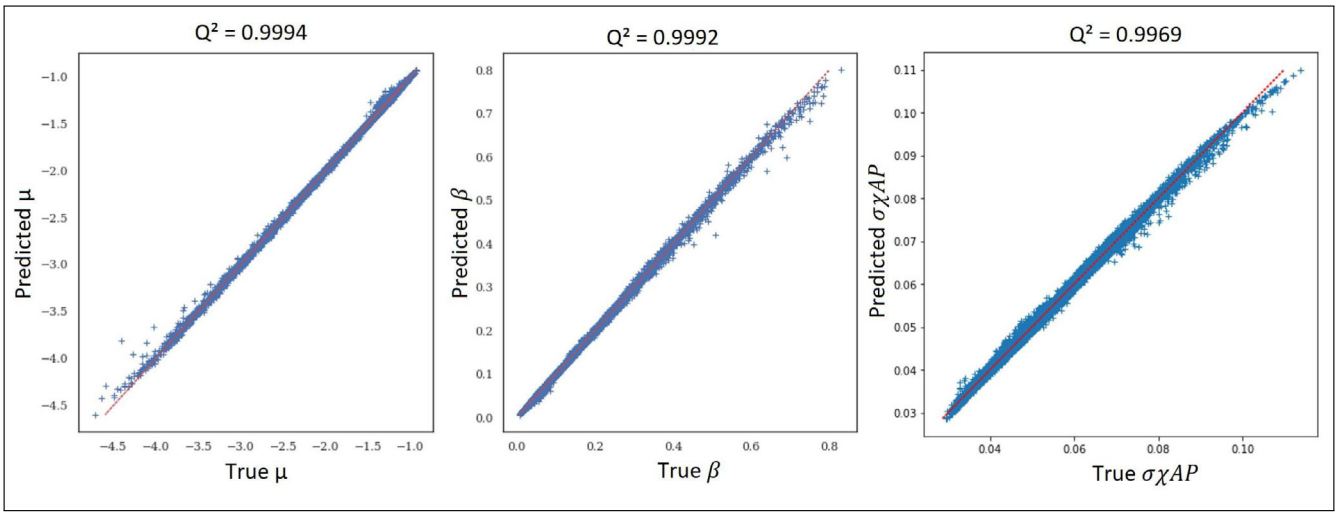

FIGURE 3 Predicted values with a Gaussian process using r_0, $\overline{h}$, and $\overline{v}$ versus real values of μ, β, and $\sigma_{\chi_{AP}}$ on the test data.

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

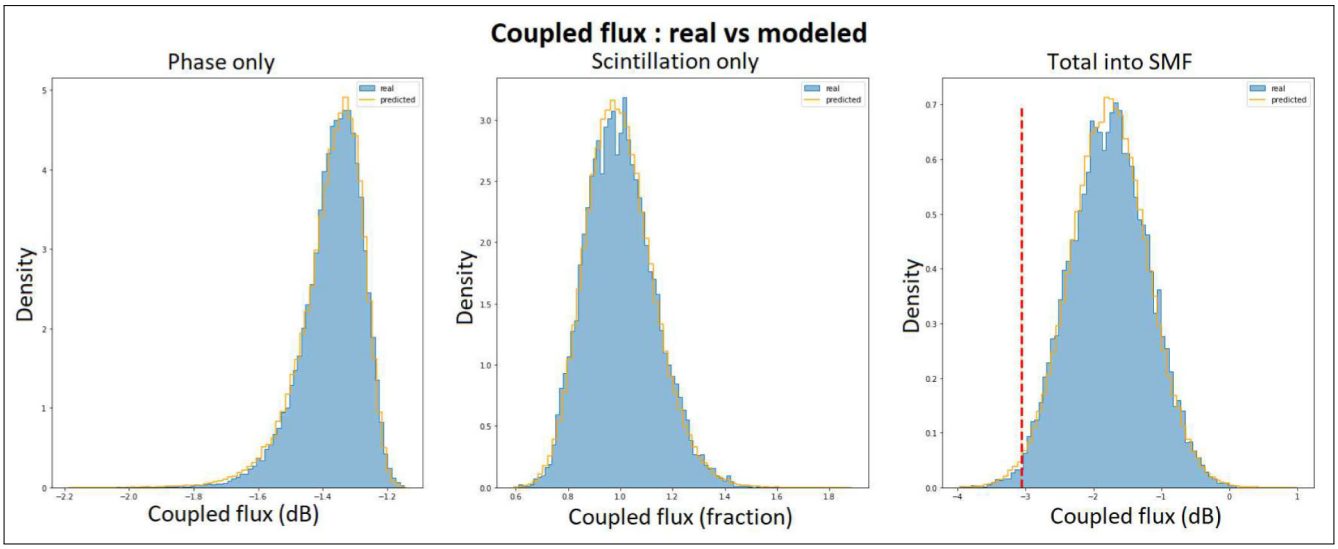

FIGURE 4 Probability density function (PDF) computed from simplified adaptive optics simulation tool (SAOST) time series and PDF computed from our metamodel with, from left to right: $L_\phi(t)$, ρ_I, and the received optical power.

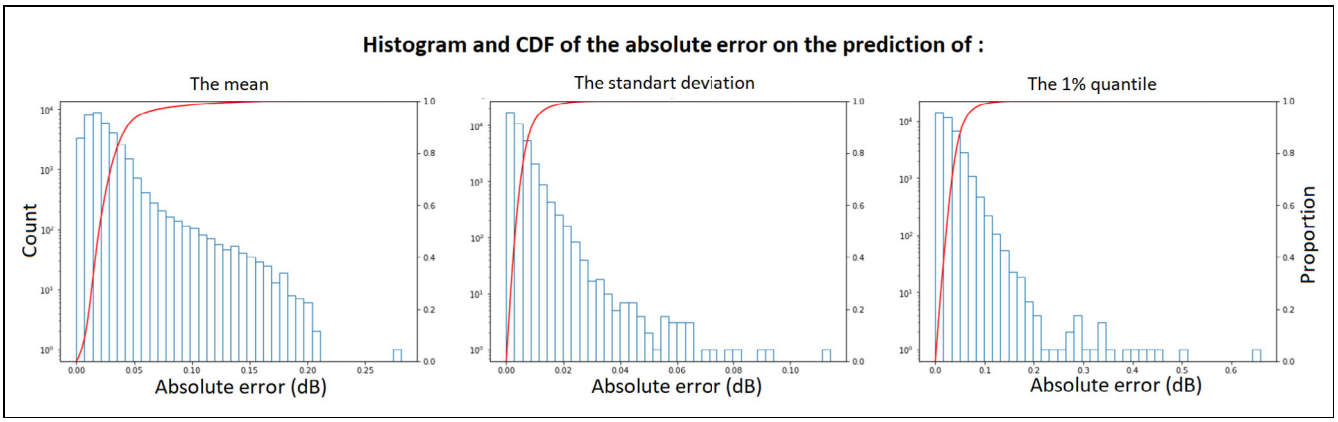

FIGURE 5 Histogram and CDF of the absolute error on prediction of the mean, the standard deviation, and the 1% quantile.

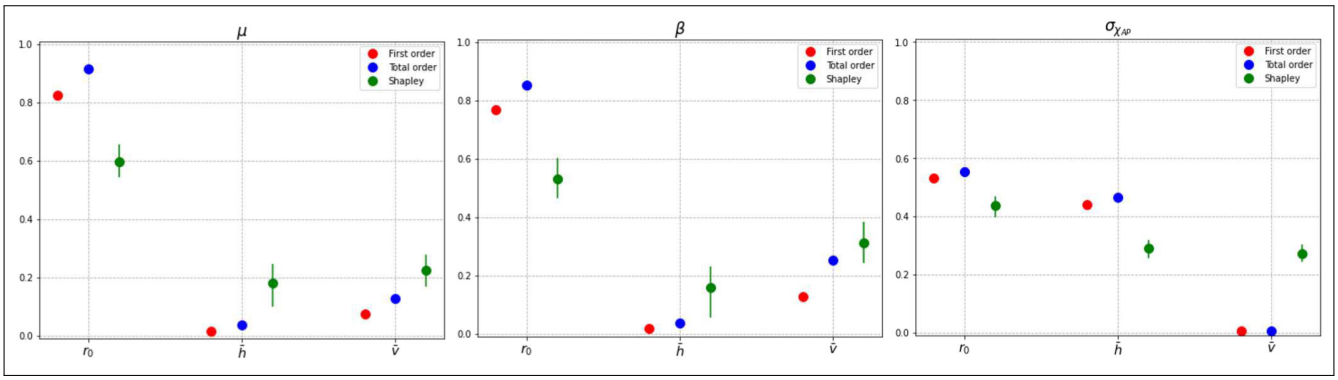

FIGURE 6 Sensitivity analysis on μ, β, and σ_{chi}. **Red:** first Sobol indices; **blue:** total Sobol indices; **green:** Shapley indices.

In Figure 6, we notice that the variability of $\overline{h}$ has little impact on μ and β. We thus can wonder if we would be able to predict the value taken by μ and β with the same accuracy if we were to fix the value of $\overline{h}$. Doing so and using the same GP structure as the one described in Sections 3.2, we obtain a prediction score of 0.9988 for μ and 0.9984 for β. The slight decrease in Q^2 is due to a bigger error on the small values of μ and

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the large values of β, that is, the cases were the residual phase error after AO correction is the highest. Even if $\overline{h}$ is not essential to adequately describe the residual phase error, the additional information it provides on the structure of the profile allows better treatment of cases of strong turbulence where the correction by AO is the worse.

Same analysis can be conducted for the prediction of $\sigma_{\chi_{AP}}$ where, according to Figure 6, the impact of $\overline{v}$ seems very low: Its total Sobol index is almost equal to zero, and its Shapley index is larger but also accounts for its correlation with the two other input variables. Removing $\overline{v}$ from the input parameters of our metamodel leads us to a prediction score of 0.9955, very close to the 0.9969 obtained with $\overline{v}$. This result was to be expected as expressions of the variance of log-amplitude averaged on a pupil that can be found in the literature are independent of the wind speed profile.[65]

4.2 | Autocorrelation assessment

The next step of our study is to be able to estimate the autocorrelation time of the ROP using the same moments. Knowledge on the temporal behavior of the ROP is fundamental as the duration of the fading in the received signal will dictate the use of numerical mitigation techniques and the latency in the telecommunication protocol.

Once again, it should be noted that analytical expressions exist to describe the auto-covariance of L_ϕ[10] and of ρ_l.[66,67] The ultimate goal is to estimate those autocovariance functions using a small number of instruments but, in order to simplify the problem, we first look at the half-correlation time of L_ϕ and ρ_l time series independently. A simplistic way to predict the autocovariance could then be to fit an exponential decay law matching the estimated half correlation time.

4.2.1 | Metamodel construction

The approach is exactly the same as for the prediction of the PDF: Half-correlation times were computed using SAOST, and we use r_0, $\overline{h}$, and $\overline{v}$ as inputs of our metamodels. Once more, best results were obtained using GP regression with a Matérn 5/2 kernel.

Figure 7 shows that the prediction using r_0, $\overline{h}$, and $\overline{v}$ is able to recreate the trend but that it lacks precision with a high variance on the distribution of the error. The chosen input moments do not account well enough for the temporal aspect of the turbulence, and it is then necessary to consider an additional measurement.

One way to recover the missing temporal information on the atmospheric layer is through measurement of the power spectral density (PSD) of the scintillation on a small pupil instrument. To simulate the measurement of such an instrument, we used the expression given in Shen et al[67] and calculated the PSD of the scintillation for a 5cm pupil at a wavelength of 1500nm. Simulation was done on 100 points spaced evenly on a log scale from $1e^{-4}$Hz up to 1000Hz.

In order to add information associated to the scintillation spectrum in our model, we wish to reduce the data dimension while keeping the maximum of information. Indeed, there is a lot of information redundancy in the 100 points used to simulate the power spectrum density, but these points have to be decimated in a nonlinear way if we want to keep the relevant information no matter the profile.

Trials were done using the cutoff frequency, slope, and magnitude of the PSD, but these added input did not result in significant improvements of the prediction score. This is due the fact that it is sometimes hard to extract and define this parameters due to multiple regimes in the PSD of the scintillation.

To extract relevant features of the PSD without analyzing each case manually, we used a convolutional autoencoder, a kind of neural network extensively used for data reduction. The underlying concept is simple; the architecture consists of two parts, an encoder and a decoder.[68] The encoder learns an encoding of the data and is validated and refined by attempting to regenerate the inputs from the latent space with the decoder. In our case, the input data are of dimension 100 and the encoded data of dimension 5, as shown on Figure 8. A value of 5 was shown to be the minimal value that enables accurate reconstruction of the PSD whatever the profiles might be; increasing the dimension of the latent space above 5 did not result in significant reconstruction improvement. In Figure 9, we can see an example of reconstruction on two randomly picked PSDs from our database. The real PSD computed from Shen et al[67] is represented in blue, while the output of the decoder applied on the encoding of the PSD in the 5 dimension latent space is in orange. We can see that most of the information is conserved in the latent space.

We built a new GP; this time with eight inputs: the three previous moments and the five encoded values of the PSD.

As anticipated, with the added temporal information, we obtained much more satisfactory results than the one described in Figure 7 (see Q^2 in Figure 10: 0.99 and 0.98 against Q^2 in Figure 7: 0.88 and 0.84)

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

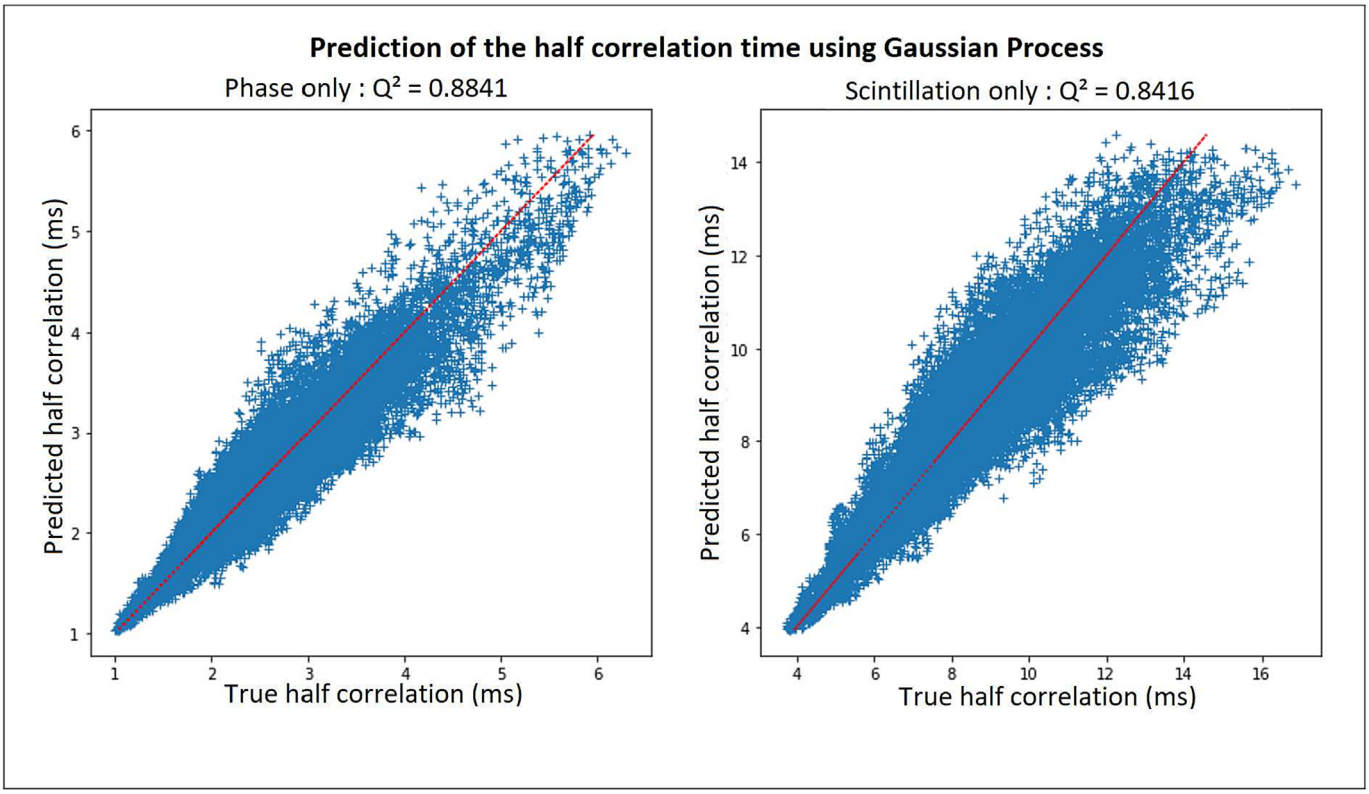

FIGURE 7 Prediction of the half-correlation time using Gaussian process (GP) on inputs r_0, $\overline{h}$, and $\overline{v}$; x-axis shows the real value and y-axis the predicted one. **On the left** are the half-correlation neglecting scintillation effects and **on the right** are neglecting phase effects.

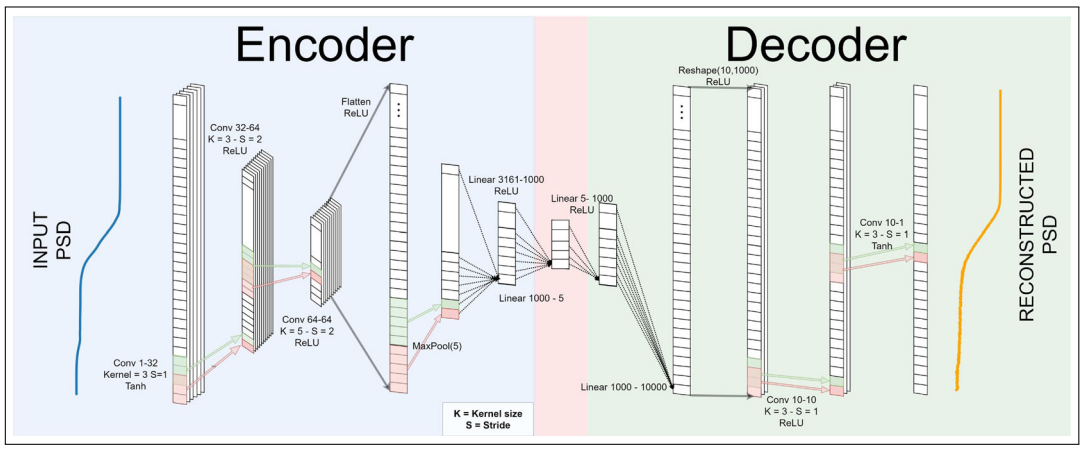

FIGURE 8 Architecture of the autoencoder used to encode the power spectral density (PSD)

4.2.2 | Sensitivity analysis

Using the same tools as the ones described in Section 4.1.2, we conducted a sensitivity analysis and calculated first and total Sobol indices as well as Shapley indices for each of the models described in Section 4.2.1.

Results for the metamodel taking only r_0, $\overline{h}$, and $\overline{v}$ as inputs can be seen in Figure 11, while sensitivity analysis results computed on the model with the PSD added in the inputs are shown in Figure 12. In both cases, we can clearly see that the most influential variable is $\overline{v}$, which is not surprising given the fact that the correlation time depends on the displacement speed of the turbulent layer.

When the five moments of the PSD are added (D1 to D5 in Figure 12), we obtain some interesting values. It can be noted that while D1 and D5 have a lot of influence on the metamodel outputs, D2 seems to have very little. It would be interesting to look further into the

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

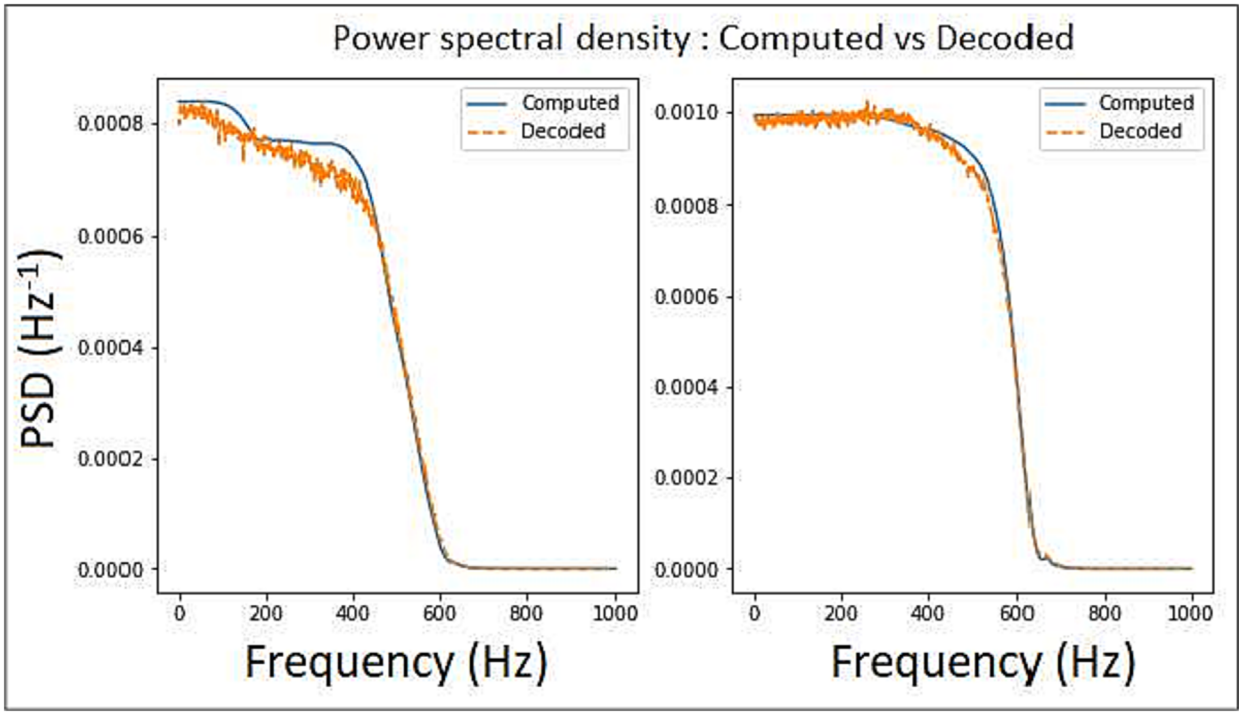

FIGURE 9 Comparison of real and reconstructed power spectral density (PSD)

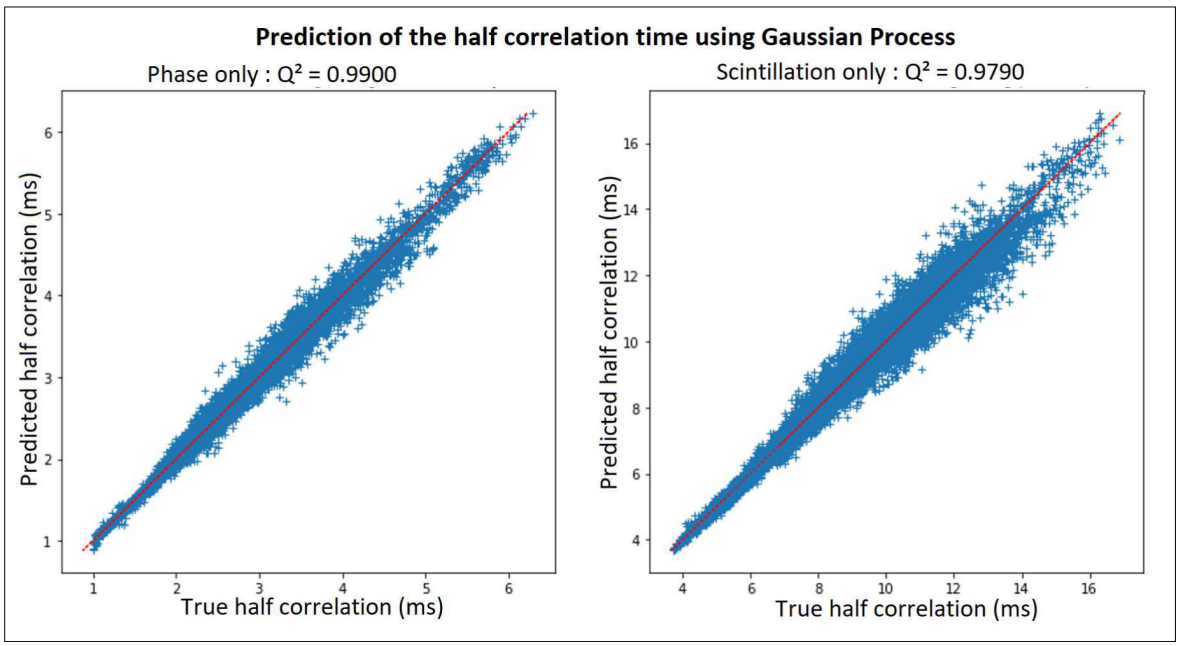

FIGURE 10 Prediction of the half-correlation time using Gaussian process on inputs r_0, $\overline{h}$, and $\overline{v}$ and information on the power spectral density (PSD) of scintillation for a 5 cm pupil. Same other parameters as in Figure 7

weights of the autoencoder to better understand what features of the PSD does each of the encoded value represent most and thus what part of the PSD function is important in the description of the demicorrelation time of the ROP. This study is not trivial and will not be conducted in the context of this article.

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

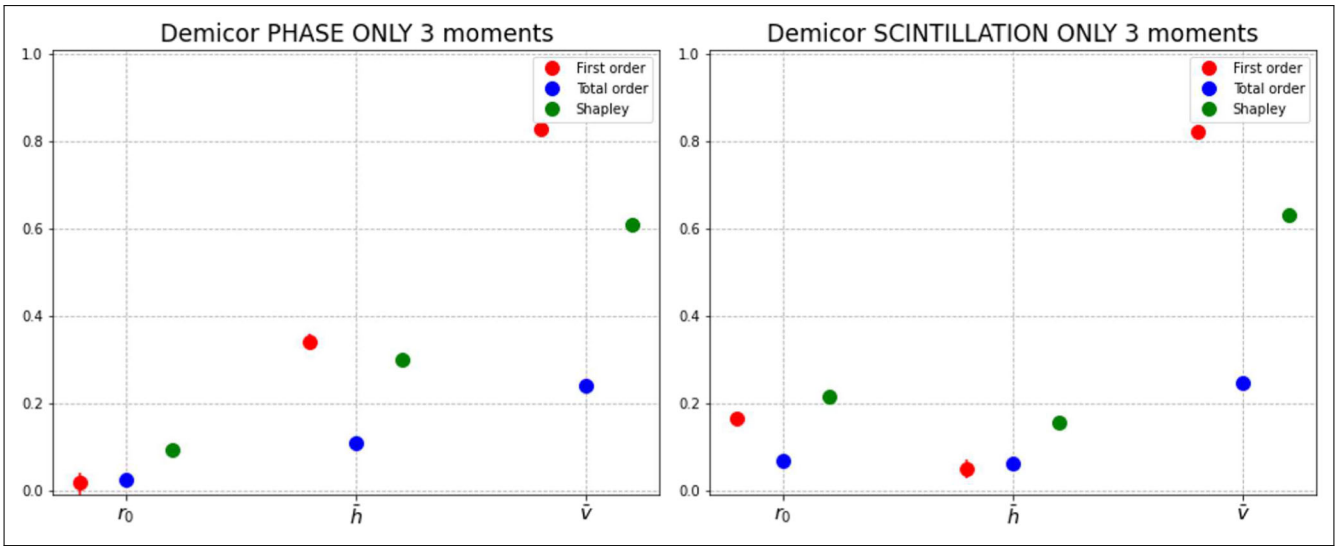

FIGURE 11 Sensitivity analysis of the model described in Figure 7. **Red:** first Sobol indices; **blue:** total Sobol indices; **green:** Shapley indices.

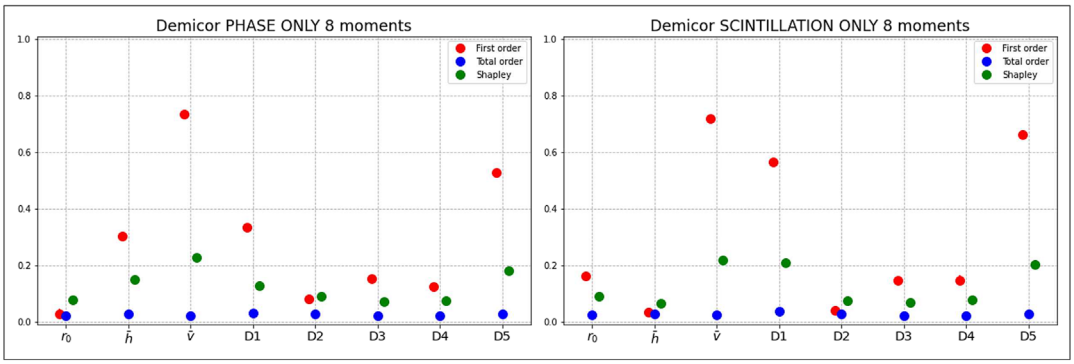

FIGURE 12 Sensitivity analysis of the model described in Figure 10. **Red:** first Sobol indices; **blue:** total Sobol indices; **green:** Shapley indices.

5 | DISCUSSIONS

The results presented in this study rely on the representativeness of the numerical model describing the influence of the propagation channel, but this model has three main limitations: (1) it assumes that we are in a Rytov regime, (2) it assumes a decorrelation of phase and amplitude disturbances, and (3) it ignores the potential influence of uncorrected static aberrations, assumes a perfect wavefront sensor and a perfect deformable mirror, and neglects the noises affecting the detection chain.

The first hypothesis, regarding compatibility with the Rytov regime, has been verified with the data used in this study, but its validity at lower altitudes or under more severe turbulence conditions could be compromised. A study should be carried out to evaluate the difference between the performance evaluation provided by our metamodel and that obtained by an end-to-end model in cases of very severe turbulence, which would deepen our understanding of the limitations of the approach.

The second hypothesis, concerning the decorrelation of phase and amplitude perturbations, is widely accepted in the community, as it is justified by the fact that the phenomena causing these perturbations are generated at different distances from the receiver, and are therefore decorrelated. The simulations mentioned in the previous paragraph would also contribute to strengthen the understanding of the limitations implied by this assumption.

Concerning the third hypothesis, on the influence of the various assumptions made to simulate the ROP, these assumptions were intentionally made to ease the understanding of limitations raised by AO before the proposed methodology being applied on data obtained by a real system. In order to validate the methodology experimentally, a proper approach could be to select experimental data so that they respect the underlying hypothesis of the physical model used to build the methodology. The possibility to extend the approach to more complex cases (where the assumptions made are no longer valid) would still have to be investigated and would probably necessitate to increase the number of parameters in the metamodel.

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

6 | CONCLUSION

We studied here the possibility to assess downlink GEO to ground optical link ROP statistics by applying a machine learning methodology. We demonstrate that, assuming several simplification hypotheses on the AO system performance model, a set of a very small number of integrated turbulence parameters appears sufficient to precisely assess ROP statistics. Moreover, exploiting the temporal power spectrum of the scintillation recorded by a small diameter receiver eases the evaluation of the ROP correlation time. An experimental validation of such a methodology could be conducted on real data, first by restricting the cases of application to those where the assumptions made to build the metamodel are satisfied. Further investigations are being conducted to extend the method to the uplink case and to evaluate the impact, in terms of performance assessment, of a deviation of the instrument aiming angle to the link line of sight.

ACKNOWLEDGMENTS

Atmospheric data were kindly provided by Durham University, whose model is commercially exploited by Miratlas.

CONFLICT OF INTEREST STATEMENT

Nicolas Védrenne : Miratlas (I); other authors declare no conflicts of interest.

REFERENCES

1. Hauschildt H, Elia C, Moeller HL, et al. HydRON: high thRoughput optical network. In: 2019 IEEE International Conference on Space Optical Systems and Applications (ICSOS); 2019:1-6. https://doi.org/10.1109/ICSOS45490.2019.8978985
2. Toyoshima M. Recent trends in space laser communications for small satellites and constellations. *J Light Technol*; 39(3):693-699. https://doi.org/10.1109/JLT.2020.3009505
3. Lacoste F, Guérin A, Laurens A, Azema G, Periard C, Grimal D. FSO ground network optimization and analysis considering the influence of clouds. In: Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP); 2011:2746-2750.
4. Wright MW, Morris JF, Kovalik JM, Andrews KS, Abrahamson MJ, Biswas A. Adaptive optics correction into single mode fiber for a low earth orbiting space to ground optical communication link using the OPALS downlink. *Opt Express*. 2015;23(26):33705-33712. https://doi.org/10.1364/OE.23.033705
5. Fischer E, Berkefeld T, Feriencik M, et al. Development, integration and test of a transportable adaptive optical ground station. In: IEEE International Conference on Space Optical Systems and Applications (ICSOS); 2015:1-6. https://doi.org/10.1109/ICSOS.2015.7425071
6. Clark Jr GC, Cain JB. *Error-Correction Coding for Digital Communications*: Springer Science & Business Media; 2013.
7. Poulenard S, Gadat B, Barthe L, Garzón-Bohórquez R. Protection schemes for optical communication between optical ground station and a satellite. In: COAT-2019-workshop (Communications and Observations through Atmospheric Turbulence: characterization and mitigation) ONERA; 2019. https://doi.org/10.34693/COAT2019-S5-002
8. Chabé J, Aristidi E, Ziad A, et al. PML: a generalized monitor of atmospheric turbulence profile with high vertical resolution. *Appl Optics*. 2020;59(25):7574. https://doi.org/10.1364/AO.384504
9. Jabet F. C-DIMM: an autonomous, outdoor and fixed seeing monitor for astronomy, atmospheric studies and free space optical communications. In: Stein K, Gladysz S, eds. *Environmental Effects on Light Propagation and Adaptive Systems IV*. Online Only, Spain: SPIE; 2021:3. https://doi.org/10.1117/12.2599211
10. Canuet L, Védrenne N, Conan J-M, Petit C, Artaud G, Rissons A, Lacan J. Statistical properties of single-mode fiber coupling of satellite-to-ground laser links partially corrected by adaptive optics. *J Opt Soc Am A*. 2018;35(1):148-162. https://doi.org/10.1364/JOSAA.35.000148
11. Aristidi E, Ziad A, Fanteï-Caujolle Y, Chabé J, Giordano C, Renaud C, Lanteri H. Monitoring daytime and nighttime optical turbulence profiles with the PML instrument. https://doi.org/10.48550/ARXIV.2002.04947; 2020.
12. Ziad A, Giordano C, Aresta A, et al. ANAtOLIA: a new mobile site-testing station for astronomy and optical communications. In: Schreiber L, Schmidt D, Vernet E, eds. *Adaptive Optics Systems VIII*, Vol. 12185: SPIE; 2022:121855Z.
13. Miratlas. https://www.miratlas.com/products_services.html
14. Wang Y, Basu S. Estimation of optical turbulence in the atmospheric surface layer from routine meteorological observations an artificial neural network approach. In: Proceedings of SPIE, Vol. 9224; 2014. https://doi.org/10.1117/12.2063168
15. Wang Y, Basu S. Using an artificial neural network approach to estimate surface-layer optical turbulence at Mauna Loa, Hawaii. *Opt Lett*. 2016;41(10):2334-2337. https://doi.org/10.1364/OL.41.002334
16. Su C, Wu X, Luo T, Wu S, Qing C. Adaptive niche-genetic algorithm based on backpropagation neural network for atmospheric turbulence forecasting. *Appl Opt*. 2020;59(12):3699-3705. https://doi.org/10.1364/AO.388959
17. Jellen C, Oakley M, Nelson C, Burkhardt J, Brownell C. Machine-learning informed macro-meteorological models for the near-maritime environment. *Appl Opt*. 2021;60(11):2938-2951. https://doi.org/10.1364/AO.416680
18. Lamprecht C, Bekhrad P, Ivanov H, Leitgeb E. Modelling the refractive index structure parameter: A ResNet approach. In: 2020 International Conference on Broadband Communications for Next Generation Networks and Multimedia Applications (CoBCom); 2020:1-4. https://doi.org/10.1109/CoBCom49975.2020.9174186
19. Grose M. Forecasting atmospheric turbulence conditions from prior environmental parameters using artificial neural networks: an ensemble study. *Ph. D. Thesis*; 2021.
20. Vorontsov AM, Vorontsov MA, Filimonov GA, Polnau E. Atmospheric turbulence study with deep machine learning of intensity scintillation patterns. *Appl Sci*. 2020;10(22):8136. https://doi.org/10.3390/app10228136
21. Milli J, Gonzalez R, Fluxa PR, et al. Nowcasting the turbulence at the paranal observatory. https://doi.org/10.48550/ARXIV.1910.13767; 2019.

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

22. Giordano C, Rafalimanana A, Ziad A, Aristidi E, Chabé J, Fanteï-Caujolle Y, Renaud C. Statistical learning as a new approach for optical turbulence forecasting. In: SPIE Astronomical Telescopes + Instrumentation, Adaptive Optics Systems VII, Vol. 11448 SPIE; 2020:114484E. https://doi.org/10.1117/12.2562316
23. Amirabadi MA, Kahaei MH, Nezamalhosseni SA. Low complexity deep learning algorithms for compensating atmospheric turbulence in the free space optical communication system. *IET Optoelectron*. 2022;16(3):93-105. https://doi.org/10.1049/ote2.12060
24. Esmail MA, Saif WS, Ragheb AM, Alshebeili SA. Free space optic channel monitoring using machine learning. *Opt Express*. 2021;29(7):10967-10981. https://doi.org/10.1364/OE.416777
25. Tóth J, Ovseník L', Turán J, Michaeli L, Márton M. Classification prediction analysis of RSSI parameter in hard switching process for FSO/RF systems. *Measurement*. 2018;116:602-610. https://doi.org/10.1016/j.measurement.2017.11.044
26. Lionis A, Peppas K, Nistazakis HE, Tsigopoulos A, Cohn K, Zagouras A. Using machine learning algorithms for accurate received optical power prediction of an FSO link over a maritime environment. *Photonics*. 2021;8(6). https://doi.org/10.3390/photonics8060212
27. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J*. 1948;27(3):379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x
28. Farid AA, Hranilovic S. Outage capacity optimization for free-space optical links with pointing errors. *J Lightwave Technol*. 2007;25(7):1702-1710. https://doi.org/10.1109/JLT.2007.899174
29. Conan J-M, Montmerle-Bonnefois A, Védrenne N, et al. Adaptive optics for GEO-feeder links: from performance analysis via reciprocity based models to experimental demonstration. In: COAT-2019 - workshop (Communications and Observations through Atmospheric Turbulence: characterization and mitigation) ONERA; 2019; Châtillon, France. https://doi.org/10.34693/COAT2019-S5-003
30. Vedrenne N, Conan J-M, Petit C, Michau V. Adaptive optics for high data rate satellite to ground laser link. In: Proc. SPIE 9739, Free-Space Laser Communication and Atmospheric Propagation XXVIII, 97390E 9739; 2016. https://doi.org/10.1117/12.2218213
31. Canuet L. Reliability of satellite-to-ground optical communication. *Ph.D. Thesis*; 2018.
32. Bonnefois AM, Velluet M-T, Cissé M, et al. Feasibility demonstration of AO pre-compensation for GEO feeder links in a relevant environment. *Opt Express*. 2022;30(26):47179-47198. https://doi.org/10.1364/OE.470705
33. Gracheva ME, Gurvich AS. Strong fluctuations in the intensity of light propagated through the atmosphere close to the earth. *Sov Radiophysics*. 1965;8(4):511-515.
34. Roddier NA. Atmospheric wavefront simulation using Zernike polynomials. *Opt Eng*. 1990;29(10):1174-1180. https://doi.org/10.1117/12.55712
35. Conan J-M, Rousset G, Madec P-Y. Wave-front temporal spectra in high-resolution imaging through turbulence. *JOSA*. 1995;12(7):1559-1570. https://doi.org/10.1364/JOSAA.12.001559
36. Tatarskii VI. The effects of the turbulent atmosphere on wave propagation; 1971.
37. Vetelino FS, Young C, Andrews L, Recolons J. Aperture averaging effects on the probability density of irradiance fluctuations in moderate-to-strong turbulence. *Appl Opt*. 2007;46(11):2099-2108. https://doi.org/10.1364/AO.46.002099
38. Wilson RW. SLODAR: measuring optical turbulence altitude with a Shack–Hartmann wavefront sensor. *Mon Not R Astron Soc*. 2002;337(1):103-108. https://doi.org/10.1046/j.1365-8711.2002.05847.x
39. Osborn J, Wilson RW, Sarazin M, et al. Optical turbulence profiling with Stereo-SCIDAR for VLT and ELT. *Mon Not R Astron Soc*. 2018;478(1):825-834. https://doi.org/10.1093/mnras/sty1070
40. Miller JE, Eaton FD, Stokes SS. Correlation of weather parameters with sodar-derived Cn2 profiles. In: Thompson WE, Merritt PH, eds. *Laser Weapons Technology II*, Vol. 4376: SPIE; 2001:116-123.
41. Greenwood DP. Bandwidth specification for adaptive optics systems* *J Opt Soc Am*. 1977;67(3):390. https://doi.org/10.1364/JOSA.67.000390
42. Good RE, Beland RR, Murphy EA, Brown JH, Dewan EM. Atmospheric models of optical turbulence. In: Modeling of the Atmosphere; 1988; Orlando, FL, United States:165. https://doi.org/10.1117/12.975626
43. Parenti RR, Sasiela RJ. Laser-guide-star systems for astronomical applications. *J Opt Soc Am*. 1994;11(1):288. https://doi.org/10.1364/JOSAA.11.000288
44. Hufnagel RE. Propagation through atmospheric turbulence. *The Infrared Handbook*: USGPO; 1974.
45. Andrews LC, Phillips RL, Wayne D, Leclerc T, Sauer P, Crabbs R, Kiriazes J. Near-ground vertical profile of refractive-index fluctuations. In: Wasiczko Thomas LM, Gilbreath GC, eds. *In Atmospheric Propagation*. Orlando, Florida, USA; 2009:732402. https://doi.org/10.1117/12.820369
46. Sadot D, Kopeika NS. Forecasting optical turbulence strength on the basis of macroscale meteorology and aerosols: models and validation. *Opt Eng*. 1992;31(2):200-212.
47. Dewan EM, Good RE, Beland RR, Brown JH. A model for CSUBN(2) (optical turbulence) profiles using radiosonde data; 1993.
48. Ruggiero FH, DeBenedictis DA. Forecasting optical turbulence from mesoscale numerical weather prediction models. In: DoD High Performance Modernization Program Users Group Conference; 2002:10-14.
49. Osborn J, Sarazin M. Atmospheric turbulence forecasting with a general circulation model for Cerro Paranal. *Mon Not R Astron Soc*. 2018;480(1):1278-1299. https://doi.org/10.1093/mnras/sty1898
50. Osborn J, Communal J-E, Jabet F. Global atmospheric turbulence forecasting for free-space optical communications; 2023.
51. Copernicus Climate Change Service. ERA5-Land hourly data from 2001 to present. Type: dataset. https://doi.org/10.24381/CDS.E2161BAC; 2019.
52. Rasmussen CE, Williams CKI. *Gaussian Processes for Machine Learning*: MIT press Cambridge, MA; 2005.
53. Tokovinin A. From differential image motion to seeing. In: Publications of the Astronomical Society of the Pacific, Vol. 114; 2002:1156-1166. https://doi.org/10.1086/342683
54. Roddier F. V the effects of atmospheric turbulence in optical astronomy. *Progress in Optics*: Elsevier; 1981:281-376.
55. Fried DL. Anisoplanatism in adaptive optics. *J Opt Soc Am*. 1982;72(1):52-61. https://doi.org/10.1364/JOSA.72.000052
56. Loos GC, Hogge CB. Turbulence of the upper atmosphere and isoplanatism. *Appl Opt*. 1979;18(15):2654-2661. https://doi.org/10.1364/AO.18.002654
57. Beckers JM. A seeing monitor for solar and other extended object observations. *Exp Astron*. 2001;12(1):1-20.
58. Roddier F, Gilli JM, Lund G. On the origin of speckle boiling and its effects in stellar speckle interferometry. *J Opt*. 1982;13(5):263-271. https://doi.org/10.1088/0150-536x/13/5/002
59. Fante RL, Electromagnetic beam propagation in turbulent media. *Proc IEEE*. 1975;63(12):1669-1692. https://doi.org/10.1109/PROC.1975.10035

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

60. Sobol IM. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Math Comput Simul*. 2001;55(1):271-280. The Second IMACS Seminar on Monte Carlo Methods. https://doi.org/10.1016/S0378-4754(00)00270-6
61. Song E, Nelson BL, Staum J. Shapley effects for global sensitivity analysis: theory and computation. *SIAM/ASA J Uncertain Quantif*. 2016;4(1):1060-1083. https://doi.org/10.1137/15M1048070
62. Iooss B, Prieur C. Shapley effects for sensitivity analysis with correlated inputs: comparisons with Sobol' indices, numerical estimation and applications. *Int J Uncertain Quantif*. 2019;9(5):493-514. https://doi.org/10.1615/Int.J.UncertaintyQuantification.2019028372
63. Iooss B, Veiga SD, Janon A, et al. sensitivity: global sensitivity analysis of model outputs. https://CRAN.R-project.org/package=sensitivity; 2021.
64. R Core Team. R: a language and environment for statistical computing. In: R Foundation for Statistical Computing; 2022; Vienna, Austria. https://www.R-project.org/
65. Ziad A, Conan R, Tokovinin A, Martin F, Borgnino J. From the grating scale monitor to the generalized seeing monitor. *Appl Opt*. 2000;39(30):5415-5425. https://doi.org/10.1364/AO.39.005415
66. Robert C, Conan J-M, Michau V, Renard J-B, Robert C, Dalaudier F. Retrieving parameters of the anisotropic refractive index fluctuations spectrum in the stratosphere from balloon-borne observations of stellar scintillation. *J Opt Soc Am A*. 2008;25(2):379-393. https://doi.org/10.1364/JOSAA.25.000379
67. Shen H, Yu L, Fan C. Temporal spectrum of atmospheric scintillation and the effects of aperture averaging and time averaging. *Opt Commun*. 2014;330:160-164. https://doi.org/10.1016/j.optcom.2014.05.039
68. Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science*. 2006;313(5786):504-507. https://doi.org/10.1126/science.1127647

AUTHOR BIOGRAPHIES

Emile Klotz is a last year PhD student in the Optics Department at ONERA (The French Aerospace Lab). He graduated with a Master's degree at *Institut d'Optique Théorique et Appliquée*. His main research interests are optics applied to space telecommunications and the integration of machine learning techniques into physical modelling.

Sidonie Lefebvre received her Ph.D. degree from Ecole Centrale Paris in 2006. She is currently working as a Research Scientist in statistics at ONERA Palaiseau (France), in the Optics and Associated Techniques Department. Her research interests are centered on the design and modeling of computer experiments, uncertainty quantification and sensitivity analysis, and target detection.

Nicolas Vedrenne joined ONERA in 2004. He defended his Phd on optical propagation and correction through strong turbulence in 2008. He developed and applied his skills on adaptive optics for different fields of applications for more than 15 years in the High Angular Resolution team: initiator of the SCO-SLIDAR method, a Cn2 profil measurement method from Shack-Hartmann data, he led several on site demonstrations of AO systems for optical links, and he was the project leader for FEEDELIO (field demonstration of precompensation by AO for ESA). N. Védrenne is since 2018 the Principal Investigator of FEELINGS, ONERA's ground station demonstrator for optical links and the head of the High Angular Resolution unit of ONERA.

Christian Musso received a PHD in Applied Mathematics in 1993 and a habilitation to supervise research in 2017 (Habilitation à Diriger des Recherches). He was hired by the French naval direction in 1986 where he supervises optimal maneuvers algorithms in Target Motion Analysis. Since 1991, he works at ONERA, The French Aerospace Lab. He is interested in signal processing applied to radar and infrared images. His research focuses on nonlinear filtering and optimization using Monte Carlo methods.

Sylvain Poulenard obtained his PhD in 2015 entitled "Laser link, an alternative to radio-frequency feeder link for high throughput satellite?". He works for Airbus Defence and Space since then and is now the system engineer responsible for the TELEO Demonstrator, an experimental optical communications payload designed to facilitate very high capacity optical feeder link communications to GEO satellite.

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Thierry Fusco is a senior research director at ONERA and the science deputy director of the Optics Department. He has 25 years of expertise in adaptive optics and post-processing for astronomy. After his PhD in 2000 on the mitigation of partial correction and anisoplanatism in AO, he was the VLT-SPHERE AO scientist (extreme AO system for direct detection and characterization of extrasolar planet) from 2004 to 2015, and he is now the AO scientist of HARMONI (which gathers both a classical and laser tomographic AO system) the first light integral field spectrograph instrument for the European ELT. Thiery Fusco received the Fabry de Gramont award from the French Otpical Society in 2009, the Scientific Excellence Award from the French Aeronautics and Astronautics Association in 2020, and the "Grand Prix Charles Defforey" from the French Science Academy in 2020. He published more than 200 papers in scientific journals, and he has an H-index larger than 50.

How to cite this article: Klotz E, Lefebvre S, Vedrenne N, Musso C, Poulenard S, Fusco T. Assessment of adaptive optics-corrected optical links statistics from integrated turbulence parameters through a Gaussian process metamodel. *Int J Satell Commun Network*. 2023;1-19. doi:10.1002/sat.1497

15420981, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/sat.1497 by Onera, Wiley Online Library on [21/09/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License