

## Article

# Statistical Analysis and Machine Learning Prediction of Fog-Caused Low-Visibility Events at A-8 Motor-Road in Spain

Sara Cornejo-Bueno <sup>1</sup>, David Casillas-Pérez <sup>1,\*</sup> , Laura Cornejo-Bueno <sup>1</sup>, Mihaela I. Chidean <sup>1</sup> , Antonio J. Caamaño <sup>1</sup> , Elena Cerro-Prada <sup>2</sup> , Carlos Casanova-Mateo <sup>2</sup> and Sancho Salcedo-Sanz <sup>3</sup> 

<sup>1</sup> Department of Signal Theory and Communications, Universidad Rey Juan Carlos, 28942 Fuenlabrada, Spain; sara.cornejo@urjc.es (S.C.-B.); laura.cornejo@urjc.es (L.C.-B.); mihaela.chidean@urjc.es (M.I.C.); antonio.caamano@urjc.es (A.J.C.)

<sup>2</sup> Department of Civil Engineering, Construction, Infrastructure and Transport, Universidad Politécnica de Madrid, 28007 Madrid, Spain; elena.cerro@upm.es (E.C.-P.); carlos.casanova@upm.es (C.C.-M.)

<sup>3</sup> Department of Signal Processing and Communications, Universidad de Alcalá, 28805 Alcalá de Henares, Spain; sancho.salcedo@uah.es

\* Correspondence: david.casillas@urjc.es

**Abstract:** This work presents a full statistical analysis and accurate prediction of low-visibility events due to fog, at the A-8 motor-road in Mondoñedo (Galicia, Spain). The present analysis covers two years of study, considering visibility time series and exogenous variables collected in the zone affected the most by extreme low-visibility events. This paper has then a two-fold objective: first, we carry out a statistical analysis for estimating the fittest probability distributions to the fog event duration, using the Maximum Likelihood method and an alternative method known as the L-moments method. This statistical study allows association of the low-visibility depth with the event duration, showing a clear relationship, which can be modeled with distributions for extremes such as Generalized Extreme Value and Generalized Pareto distributions. Second, we apply a neural network approach, trained by means of the ELM (Extreme Learning Machine) algorithm, to predict the occurrence of low-visibility events due to fog, from atmospheric predictive variables. This study provides a full characterization of fog events at this motor-road, in which *orographic* fog is predominant, causing important traffic problems during all year. We also show how the ELM approach is able to obtain highly accurate low-visibility events predictions, with a Pearson correlation coefficient of 0.8, within a half-hour time horizon, enough to initialize some protocols aiming at reducing the impact of these extreme events in the traffic of the A-8 motor road.

**Keywords:** low-visibility events; orographic and hill-fogs; extreme learning machines; prediction problems; machine learning algorithms



**Citation:** Cornejo-Bueno, S.; Casillas-Pérez, D.; Cornejo-Bueno, L.; Chidean, M.I.; Caamaño, A.J.; Cerro-Prada, E.; Casanova-Mateo, C.; Salcedo-Sanz, S. Statistical Analysis and Machine Learning Prediction of Fog-Caused Low-Visibility Events at A-8 Motor-Road in Spain. *Atmosphere* **2021**, *12*, 679. <https://doi.org/10.3390/atmos12060679>

Academic Editors: Sándor Baran and Annette Möller

Received: 12 April 2021

Accepted: 23 May 2021

Published: 26 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

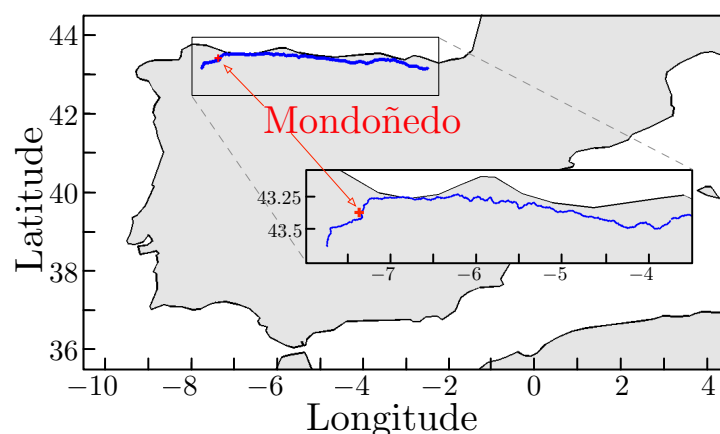
Low-visibility events are extreme atmospheric situations which deeply affect transport and transportation facilities [1–3]. They cause a large number of deaths in developed countries, associated with traffic accidents [4,5]. It is therefore a serious hazard to vehicles when visibility varies over short distances or periods of time at so-called fog “black spots” or fog-walls [6]. On the other hand, low-visibility events reduce the transport capacity on affected roads, leading to closures in extreme cases, with the consequent economic cost [7].

These extreme situations are recurrent in an important motor-road in Spain, the A-8 at Mondoñedo (Galicia, Spain). The A-8 motor-road is the largest and most important highway running along northern Spain, all along the Cantabrian coast. It runs from the Basque Country to the province of Lugo in Galicia, as can be observed in Figure 1. The geographical situation and the weather conditions of the place favor the occurrence of low-visibility events due to orographic or hill fogs at specific points (specially at “Alto de O Fiouco” area). This issue remains as an unresolved problem since the construction of

the motor-road some years ago, with an average of over 700 h a year of closure due to fog events. Solving this problem has become a priority for the local and national government, not only to reduce the number of traffic crashes associated with low-visibility events in the motor-road, but also to reduce the economic impact in the zone. In fact, there are studies such as [8], which have shown differences in economic growth after the construction of important traffic infrastructures such as motor-roads, i.e., these infrastructures have an impact on population dynamics and business performance. In this case, the economic impact of the A-8 motor-road is deeply affected by the issues related to fog events. As part of the solution to this problem, it is important the characterization of fog events, both in terms of their statistical physics, and also in terms of the meteorological causes of fog formation. It is also important the development of robust approaches to carry out an accurate prediction of the phenomenon, within a short-time prediction horizon. There are different previous works dealing with statistical characterization of different fog properties, such as its short-term and long-term persistence [9–11], dynamic processes [12], onset and duration [13] or general statistical characterization of fog for specific zones [14,15], and some other studies dealing with meteorological causes affecting fog formation [16–18]. In turn, regarding previous works dealing with prediction algorithms for fog events, very different techniques have been developed and applied. Numerical weather prediction is one of the most widely used approaches [19–22]. However, due to the local nature of this type of phenomenon, forecasting low visibility events by means of numerical weather prediction is a complex process. This is partly because fog formation is very sensitive to small-scale variations in atmospheric variables (changes in wind or in atmospheric stability) and many current models do not capture such spatial resolution. Moreover, the prediction of fog events is very sensitive to initial conditions in the numerical methods [23]. Alternative approaches involve statistical methods to predict fog events. One of the first attempts introduced the use of linear regression for marine fog events prediction [24]. In the last decade, Machine Learning (ML) techniques have been successfully applied to the prediction of low-visibility extreme events related to fog. Examples include neural networks [1,2,25], Bayesian decision approaches [26], different ML regression techniques such as Support Vector Regression or Extreme Learning Machine (ELM) algorithms for regression [27–31], evolutionary neural approaches [32] or ordinal regression techniques [33]. Note that many of these previous works are related to fog prediction in transportation facilities such as airports.

In this work we carry out a detailed study on real visibility data collected on the A-8 motor-road at Mondoñedo, Galicia, Spain. It includes the statistical characterization of the fog event, and also its prediction by using an ELM algorithm. First, we deal with the statistical characterization of the fog events at this zone. To this end, we try to obtain the fittest probability distributions to the low-visibility events available, considering its duration. We compare two methods for this task: the Maximum Likelihood and the L-moments method. For this purpose, we have tried distributions appropriate for extreme events, since the most general approach to the study of rare extreme events is based on the Extreme Value Theory (EVT) [34], especially in cases with very few samples available [35]. Second, we propose an accurate prediction algorithm for low-visibility events in the A-8 motor-road based on ELMs [36] and exogenous atmospheric variables collected in a measuring station at the motor-road. We will show that the ELM algorithm provides an excellent performance in terms of error metrics, with a low computational cost.

The rest of this paper has been structured as follows: next section presents the data description and the variables considered in the analysis and prediction of fog events at A-8 motor-road. In this section we also describe the considered distributions and the main concepts of the ELM network. Section 3 presents the experimental part of the work. First we carry out a statistical analysis of low-visibility events' duration using probability distribution fitting. We then show the performance of the ELM in the short-term prediction of these low-visibility events. Finally, Section 4 closes the paper with some concluding remarks about this study.



**Figure 1.** Location of the A-8 motor-road (in blue), and the situation of the measuring station in Mondoñedo (+ symbol in red), in Galicia, Spain. The zoomed image precisely determines the coordinates of Mondoñedo and the A8 motor-road.

## 2. Data and Methods

### 2.1. Data Description

We consider visibility data from a weather station located at Mondoñedo (43.3841 N, 7.3692 W), Galicia, Spain (Figure 1). In this area the formation of orographic or hill fog events is quite common. It directly impacts the visibility in the motor-road on deep episodes leading to the closure of the motor-road. The weather station is equipped with a Biral WS-100 visibility sensor. In this study we start considering as low-visibility events all visibility data below 2000 m, since this is the limit of visibility value provided by the sensor. Subsequently, the same analysis is performed in Section 3.1.2 for different (more restrictive) thresholds in the definition of low-visibility events: 600, 300 and 50 m. In these cases, we consider as low-visibility event any measurement under these different thresholds. The time series considered encompasses 23 months of data (from 1 January 2018 to 30 November 2019). In the prediction process, we evaluated the occurrence of low-visibility conditions by considering exogenous meteorological variables (also registered by the weather station), in order to take into account the atmospheric state. All the variables considered in this paper are summarized in Table 1.

**Table 1.** Description of the prediction (input) variables in the database for fog events description and forecasting at Mondoñedo, Galicia, Spain.

Variable	Units
Accumulated precipitation	mm/24 h
Salinity	%
Visibility (target)	m
Relative Humidity	%
Air temperature	°C
Floor temperature	°C
Dew temperature	°C
Global solar radiation	W/m <sup>2</sup>
Wind speed	km/h
Wind direction	Degrees
Atmospheric pressure	hPa

The temporal resolution of all the variables and target is 5 min.

### 2.2. Methods for Statistical Analysis of Fog Events Duration

In this work we study the behavior of the fog events attending to their duration in each season and during the two years of the study, 2018 and 2019. Estimating the

probability distributions of the fog duration in different seasons guarantees a certain level of stationarity, in the sense that we can expect that the forcing of wind, radiation, average temperatures, etc., are similar in each season (also the orographic conditions are similar).

In this respect, our main goal consists of identifying the occurrence of fog events with some theoretical probability distributions, and how they change depending on the season and thus on the physical processes that take place in each of them. Table 2 summarizes the theoretical probability distributions used in the study together with their Cumulative Distribution Function (CDF).

**Table 2.** Theoretical distributions used in the fog events duration analysis carried out. The non-negative variable  $x$  refers to the duration of fog events.

Distribution	CDF
Exponential (EXP)	$F(x; \lambda) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$
Logistic (LOG)	$F(x; \mu, s) = \frac{1}{1 + e^{-(x-\mu)/s}}$
Normal (NRM)	$F(x; \mu, \sigma) = \Phi\left(\frac{x-\mu}{\sigma}\right)$
Generalized Pareto (GPa)	$F(x; \xi) = \begin{cases} 1 - (1 + \xi x)^{-1/\xi} & \text{for } \xi \neq 0 \\ 1 - e^{-x} & \text{for } \xi = 0 \end{cases}$
Generalized Extreme Value (GEV) <sup>1</sup>	$F(x; \mu, \sigma, \xi) = \exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \quad \xi \neq 0$
Log-Normal (LN)	$F(x; \mu, \sigma) = \Phi\left(\frac{(\ln x) - \mu}{\sigma}\right)$
Gamma (GAM)	$F(x; \alpha, \beta) = \frac{\gamma(\alpha, \beta x)}{\Gamma(\alpha)}$
Extreme Value <sup>2</sup> (EV)	$F(x; \mu, \sigma) = \exp\left\{-\exp\left\{-\frac{x-\mu}{\sigma}\right\}\right\}$
Log-Logistic (LLG)	$F(x; \alpha, \beta) = \frac{1}{1 + (x/\alpha)^{-\beta}}$
Stable (STA)	Does not have an explicit CDF <sup>3</sup>

<sup>1</sup> We present here the most frequent case of the GEV distribution where  $\xi \neq 0$  and  $\xi\left(\frac{x-\mu}{\sigma}\right) > -1$ . If  $\xi = 0$  the distribution converges to an Extreme Value distribution, see Table 2. For  $\xi\left(\frac{x-\mu}{\sigma}\right) \leq -1$  and  $\xi > 0$  or  $\xi < 0$ , the GEV distributions is 0 and 1 respectively. <sup>2</sup> Additionally known as Gumbel distribution. <sup>3</sup> The CDF for the STA distribution does not have an explicit formula in general. Only for some suitable parameters does it have a closed-form expression. However, the CDF can be described in terms of its characteristic function. See reference [37] for more details. The  $\lambda, \mu, \sigma, s, \xi, \alpha, \beta$ , are the parameters of the distributions;  $\Phi$  is the standard NRM distribution function.  $\Gamma$  is the gamma function evaluated in  $\alpha$ , and  $\gamma$  is the lower incomplete gamma function. More details about  $\Gamma$  and  $\gamma$  functions can be read in the reference [38] if the reader is interested.

We chose the most used distributions which better represent the duration events (the *life* of the events), which are commonly named as lifetime distributions. All these theoretic distributions are described by a few parameters, usually 2 parameters, except for the stable distribution which needs 4. We include among them light-tail distributions, such as the exponential (EXP), the normal (NRM) or the gamma (GAM) distributions; and long-tail (heavy-tail) distributions, such as the Log-Normal (LN) and Generalized Pareto (GPa) distributions. Some of them are more general distributions, such as the Stable (STA) or the Generalized Extreme Value (GEV) distributions. Almost all of the proposed probability density functions are positive semi-definite as the low-visibility event duration variable is non-negative. Therefore, some symmetric probability distributions may be considered by left-truncating their domains, such as the normal distributions. If the goals of the present work were to obtain a phenomenological fitting to seasonal probability distributions with no relation whatsoever to a given random variable, it would be perfectly acceptable to truncate PDFs defined in the line of real numbers and accordingly adjust for the modified area. However, since we are searching for meaningful random variable distributions that we can relate to physical processes (heavy-tailed, as we can deduce from the results), no amount of truncation can adjust for the shape of the tail relative to the main mode of the distribution. The poor results obtained by the normal distribution in the analysis is a red flag referring to the potential improvement that it could be obtained by using truncated symmetrical probability distributions.



Extreme distributions were mainly used, since they are more suitable to determine the correct distribution from a finite set of samples [35]. Furthermore, they detect the maximum values better than other types of probability distributions. Due to the fact that our data are distributed forming long tails at first sight, it is expected that the use of these distributions is appropriate (as we see in Section 3). Furthermore, because extreme values are special cases of order statistics and, conversely, problems that involves order statistics can be solved by the EVT [39].

For estimating the parameters associated with the mentioned theoretical distributions (and summarized in Table 3), two different approaches were carried out: the Maximum Likelihood criterion [40,41] and another more recent approach in terms of application, the L-moments approach [42].

**Table 3.** Range of the distribution parameters.

Distribution	Range of Parameters
EXP	$\lambda > 0$
LOG	$\mu \in \mathbb{R}, s > 0$
NRM	$\mu \in \mathbb{R}, \sigma^2 > 0$
GPa	$\zeta \in (-\infty, \infty)$
GEV	$\mu, \zeta \in \mathbb{R}, \sigma^2 > 0$
LN	$\begin{cases} \mu \in (-\infty, +\infty) \\ \sigma > 0 \end{cases}$
GAM	$\alpha, \beta > 0$
EV	$\mu, \in \mathbb{R}, \sigma^2 > 0$
LLG	$\alpha, \beta > 0$
STA	$\begin{cases} \alpha \in (0, 2] - \text{stability parameter} \\ \beta \in [-1, 1] - \text{skewness parameter} \\ c \in (0, \infty) - \text{scale parameter} \\ \mu \in (-\infty, \infty) - \text{location parameter} \end{cases}$

The first method mentioned is a common method for parameter estimation and model fitting. It allows estimation of the parameters of a probabilistic model, or the coefficients of a mathematical model, so that they are the most probable from the data obtained. Mathematically, the goal of the Maximum Likelihood is to find the value of the parameters  $\theta$  which maximize the likelihood function  $L_n(\theta, \mathbf{x})$  of an  $n$ -dimensional observed sample  $\mathbf{x} = (x_1, \dots, x_n)$  over the parameter space  $\Theta$ :

$$\theta_{ML} = \arg \max_{\theta \in \Theta} L_n(\theta, \mathbf{x}) \tag{1}$$

For some distributions with 1-dimensional parametric spaces, the maximum likelihood estimator has an analytic expression, such as for the  $\lambda$  parameter of the EXP distribution which results in the sample mean. For distributions with  $n$ -dimensional parametric spaces, they must be numerically computed [43].

The L-moments approach is an alternative theory to traditional moments theory. L-moments can be estimated by linear combinations of order statistics, i.e., by L-statistics [44]. That is, you can obtain moments from the sorted data. The advantages over using traditional moments is that L-moments are able to characterize a wider range of distributions and are more robust to the presence of outliers in the data. In addition, parameters estimations by L-moments are sometimes more accurate in small number of samples than estimations by the Maximum Likelihood criterion [42]. For this reason it is not necessary to handle a large amount of data in order to make inference. Briefly, the L-moments approach estimates the parameters solving the  $n$ -dimensional generally non-linear system forms by equalizing the population L-moments  $\{\lambda_i(\theta)\}$  associated with the population distribution  $F(x)$  (see Table 2). Those  $F(x)$  are functions of the parameters  $\theta$  to be estimated, with the sample L-moments  $\{l_i(\mathbf{x})\}$ , which uniquely depend on the sample data:

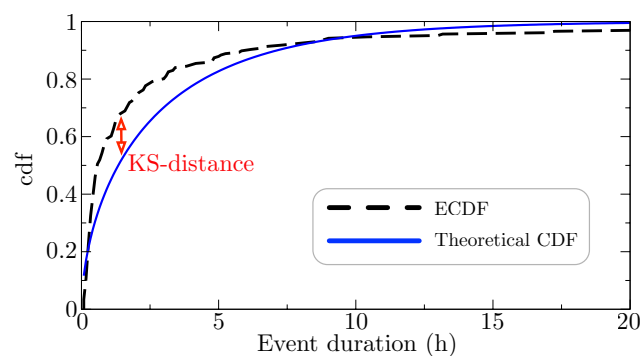
$$\lambda_i(\theta) = l_i(\mathbf{x}), \quad 1 \leq i \leq N \tag{2}$$

We require as many equations ( $N$ ) as the number of parameters to be determined to complete rank (depending of the theoretical distribution  $F(x)$  see Table 2). General expressions for population  $\{\lambda_i\}$  and sample  $\{l_i\}$  L-moment can be found in [44]. For all the evaluated theoretical distributions, there are explicit inverse functions which avoid numerical calculation of the system (2), see the work [44].

For evaluating the goodness of the distribution estimation, there exist multiple functional metrics which analyze the similarity between the estimated distributions and the data distribution. In this work, we employ the Kolmogorov–Smirnov (KS) statistic  $d_{ks}$  [45], which has the following expression:

$$d_{ks} = \sup_x |F_n(x) - F(x)| \quad (3)$$

This function measures the maximum (or supremum) distance between the cumulative distributions, the CDF  $F(x)$  and the Empirical CDF (ECDF)  $F_n(x)$ , see Figure 2. Note that the best estimation of  $F_n(x)$  will minimize Equation (3).

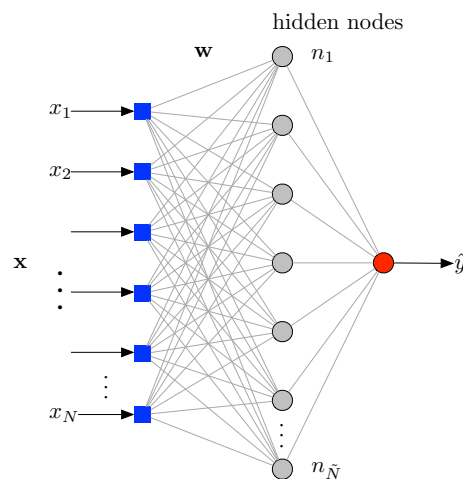


**Figure 2.** KS-distance diagram. ECDF is represented with a dotted black curve. CDF is represented with a continuous blue curve. The supremum of the distances between both curves is the KS-distance.

### 2.3. ELM for Accurate Prediction of Fog Events

Neural networks are information processing algorithms of the Artificial Intelligence family, able to efficiently solve hard problems of classification and regression from data, among other applications. The operation of feed-forward neural networks (Figure 3) starts by introducing a series of input variables (features) to the system, with known associated labels (classification) or outputs (regression). The internal layers carry out guided (by means of a training process) non-linear combinations of the inputs with their associated weights. There are several well-known algorithms for carrying out this training of the system, such as the back-propagation method [46], the Levenberg–Marquardt algorithm [47] or the ELM-type training method, which is the one applied in this work. As a result of this learning phase, we obtain a final prediction model that can be evaluated (tested) with new samples, to obtain the final predicted output. The comparison of this prediction with the real or actual output allows us quantifying the goodness of the model by means of an error measurement.

Specifically, the ELM [36] is a fast training method mainly used for feed-forward multi-layer perceptron structures, see Figure 3. In the ELM algorithm the network weights of the first layer are randomly set, usually using an uniform probability distribution. Then, it establishes the output matrix of the hidden layer and computes the Moore–Penrose pseudo-inverse of this matrix. The optimal weights of the output layer are directly obtained by multiplying the computed pseudo-inverse matrix with the target, that is, the weights of the output layer which fit best with the objective values (see [48] for details). This method obtains competitive results with respect to other classical training methods, and the training computation efficiency overcomes multi-layer perceptrons, or even Support Vector Machine algorithms [48].



**Figure 3.** Multi-layer perceptron structure considered in the ELM algorithm. It has  $N$  different inputs and the hidden layer is composed by  $\tilde{N}$  neurons.

Mathematically, the ELM algorithm considers a training set of input–output pairs  $\{(\mathbf{x}_i, y_i) \mid \mathbf{x}_i \in \mathbb{R}^N, y_i \in \mathbb{R}\}_{i=1}^n$ . Each input–output pair  $(\mathbf{x}_i, y_i)$  is a realization of the input variable vector  $\mathbf{x}$  and the output scalar variable  $y$ , which are related by a function we want to learn. In this work, the output  $y$  will be the visibility expressed in meters and the inputs  $\mathbf{x}$  will be a collection of atmospheric variables, specifically the shown in Table 1. The main goal is to fit the weights  $\{\beta_i\}$  of the output layer which multiply each output of the  $\tilde{N}$  hidden nodes. It follows the next steps:

1. Randomly assign input weights  $\mathbf{w}_i$  and the bias  $b_i$ , where  $i = 1, \dots, \tilde{N}$ , using a uniform probability distribution in  $[-1, 1]$ .
2. Calculate the hidden-layer output matrix  $\mathbf{H}$ , defined as follows:

$$\mathbf{H} = \begin{bmatrix} g(\mathbf{w}_1 \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}} \mathbf{x}_1 + b_{\tilde{N}}) \\ \vdots & \cdots & \vdots \\ g(\mathbf{w}_1 \mathbf{x}_n + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}} \mathbf{x}_n + b_{\tilde{N}}) \end{bmatrix}_{l \times \tilde{N}} \quad (4)$$

where  $g(x)$  is an activation function.

3. Finally, calculate the output weight vector  $\beta$  as follows:

$$\beta = \mathbf{H}^\dagger \mathbf{T}, \quad (5)$$

where  $\mathbf{H}^\dagger$  is the Moore–Penrose inverse of the matrix  $\mathbf{H}$  [36], and  $\mathbf{T}$  is the training output vector,  $\mathbf{T} = [y_1, \dots, y_n]^T$ .

Note that the number of hidden nodes  $\tilde{N}$  is a free parameter to be set before the training of the ELM algorithm, and must be estimated for obtaining good results by scanning a range of  $\tilde{N}$ . In this paper, we use the ELM implemented in Matlab by G. B. Huang, freely available at [49].

### 3. Experiments and Results

#### 3.1. Statistical Characterization of Fog Events Duration at A-8 Motor-Road

The statistical characterization of low visibility events associated with fog at the A-8 motor-road is discussed in this experimental subsection. The KS-distances (3) between the considered CDFs (see Table 2) and the ECDF for each year and season have been calculated taking into account the two approaches described in Section 2.2, i.e., Maximum Likelihood and L-moments methods. A seasonal analysis has been taken into account, in order to discuss whether the statistical characterization of fog events is independent (or not) from

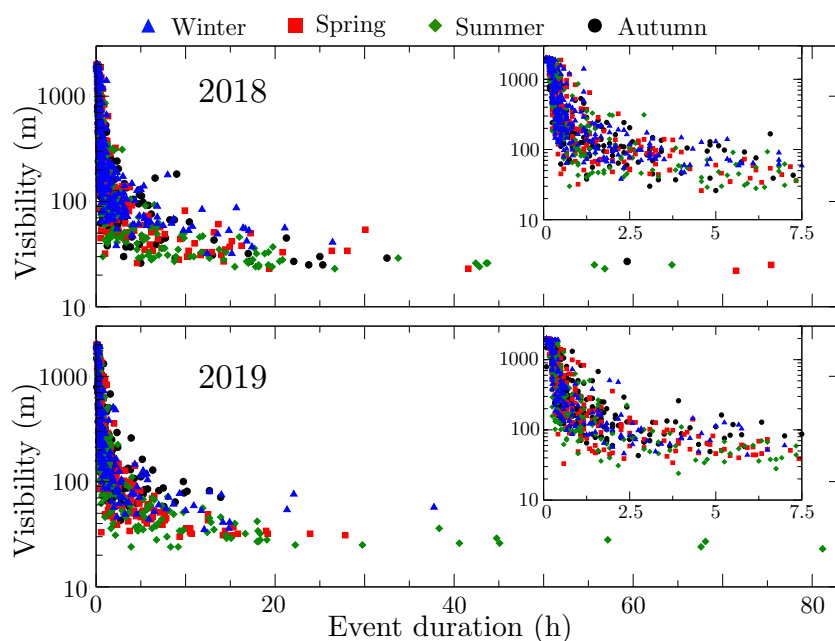
the season of the year. First, Table 4 shows the statistics of the low-visibility events at Mondoñedo in the years of the study (2018 and 2019).

**Table 4.** Low-visibility events statistics at Mondoñedo station for years 2018 and 2019, and for the 2000 m threshold.

	Fog Events Average Duration (Minutes)		# of Fog Events	
	2018	2019	2018	2019
Winter	122.69	149.22	285	168
Spring	201.42	144.02	246	246
Summer	317.63	346.98	224	191
Autumn	161.23	81.21	206	276

As can be seen, the number of low-visibility events (<2000 m) by season in both years is quite high, over 200 episodes or quite close to 200 in all cases.

In general, the number of fog events does not seem to have a dependence with the season, as in the case of radiative-type fogs, which are associated with cool/cold periods. In this case, since we have here a type of orographic fog, the number of low-visibility events is very similar in all seasons. In fact, as can be seen in Table 4, the longest average duration of the events is found in summer, with durations about 300 min in both years of study. The shortest durations appear in autumn with 81.21 min of average duration in 2019, or winter with 122.69 min in 2018. Figure 4 shows the different fog events occurred in years 2018 and 2019 at Mondoñedo and their minimal visibility associated value. Observe that those events of longer duration are associated with the lowest visibility minimum value, i.e., denser fog events take longer to dissipate.



**Figure 4.** Representation of minimum visibility vs. event duration for 2018 and 2019, and for the 2000 m threshold. Low visibility events from the different seasons are distinguished by different markers and colors. For each year, the inset in the upper right corner represents a zoom in the horizontal axis.

We proceed now with the analysis of the probability distributions for fog events duration, by means of minimizing the KS-distance for the two methods of evaluation considered: Maximum Likelihood and L-moments. We start with the Maximum Likelihood case. Table 5 shows the obtained numerical KS-distance using the Maximum Likelihood method. Moreover, Figure 5 illustrates the estimated distributions where the y-axis is plotted logarithmically scaled. In this case, ten distribution functions described in Table 2

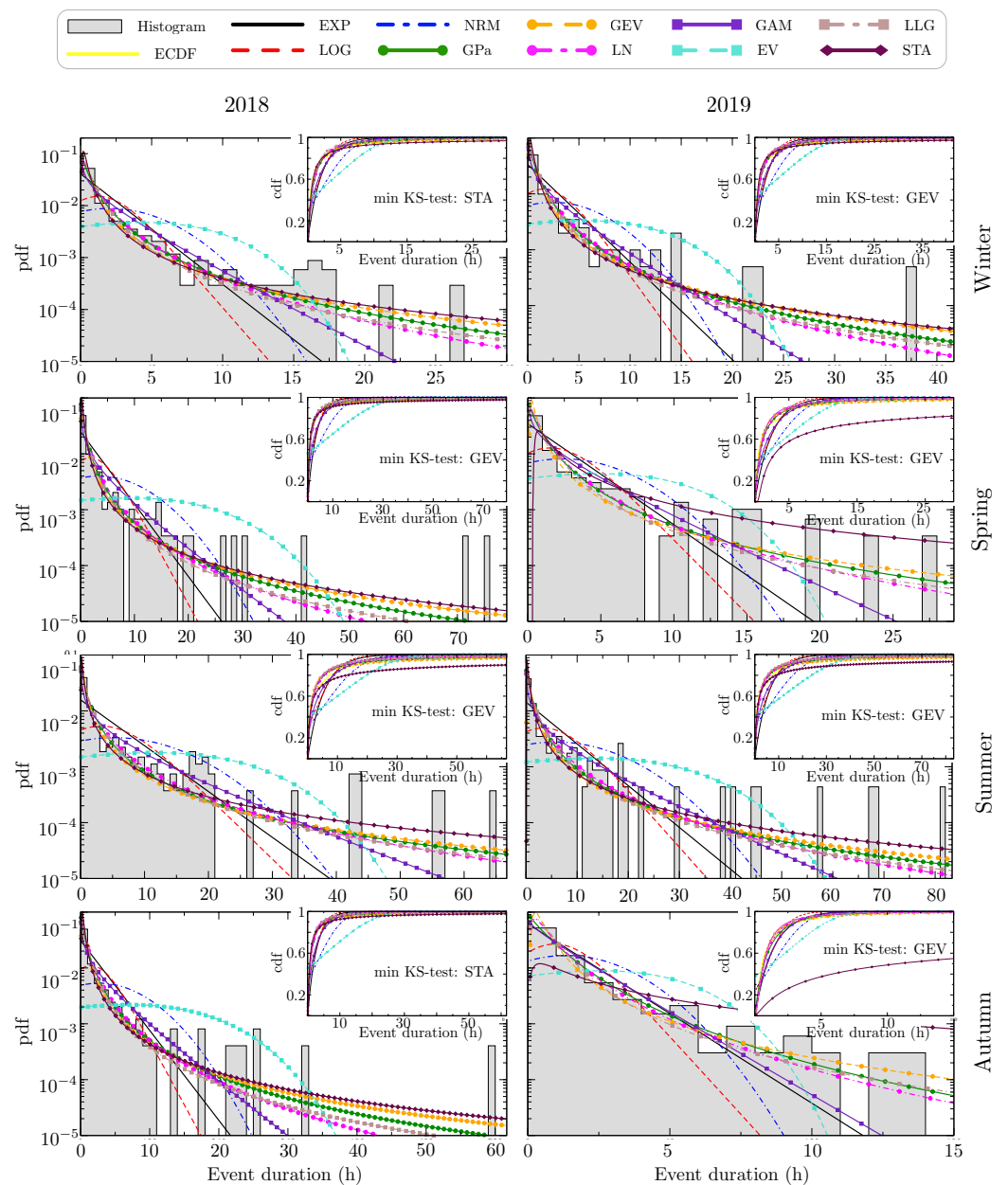
were analyzed. We show which distributions best characterize the fog events duration, based on this statistical and seasonal analysis. Specifically we show the KS-distance, which provides a quantitative reference for the best distribution. According to the results, the extremes are significant; i.e., there are many long-term events, or in other words, extreme events of low visibility, and we can see that best results in terms of KS-distance are obtained with the heavy-tail distributions. There are three distributions which fit better than the rest the duration of fog events at Mondoñedo along the seasons: These are the GEV, GPa and STA distributions. These three distributions have the shortest KS-distance and a stationary behavior through the time period. The GEV has a KS-distance around 0.07 through the seasons in both years of study, around 0.09 for the GPa, and 0.06 for the STA in 2018, but with worse values in 2019. LLG also fits well to the data distribution among all seasons around 0.1 in 2018 and 0.08 in 2019, even better than STA this last year. This can also be seen more visually in Figure 5, where heavy-tailed distributions such as GEV (in orange) or STA (in burgundy) best fit the data. The same is not true for light-tailed distributions such as EXP (in black) or NRM (in blue) and also for Logistic (LOG) (the red curve). In fact the EXP and LOG, are straight lines far from the origin in logarithmic representation, so that they do not fit very well, or NRM and EV (in cyan) which have a concave shape in this representation. Especially when we find more extreme events in the data since, as previously mentioned, those distributions are characterized by a rapid decrease in the probability of generating extreme values.

If we look at Table 5, we see that, in most seasons, in both years of study, the shortest KS-distance on average through seasons is obtained by the GEV distribution. It is known that this distribution is well suited for estimating the maximum of samples of size  $n$ , from sufficiently long sequences of independent and identically distributed random variables [50]. On the other hand, stable type distributions explain more adequately the extreme or rare phenomena, since they usually explain observations with extreme values and skewness. This denotes the presence of heavy tails [51]. This justifies the inclusion of this type of distribution in the study. Note that these distributions are a more efficient alternative to analyze high volatility phenomena due to their capacity to generate extreme values [51]. Finally, the GPa distribution also plays an important role in the EVT, and it is very common in the study of extreme events related to hydrological issues [52,53]. Its adjustment in our results shows a quite stationary behavior (in both methods) in spite of not showing the smallest KS-distance. This is not the case with the short tail distributions used in the study, i.e., EXP, LOG, NRM and EV. These distributions adapt worse to the fog events data than heavy-tail distributions, as they are characterized by a rapid decrease in the probability of generating extreme values, with KS-distance values around 0.3, or even 0.4 for the EV.

**Table 5.** KS-distances results for the Maximum Likelihood, and for the 2000 m threshold.

	2018				2019			
	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn
EXP	0.294	0.342	0.340	0.311	0.289	0.282	0.343	0.172
LOG	0.302	0.331	0.309	0.321	0.302	0.291	0.313	0.247
NRM	0.301	0.347	0.302	0.337	0.305	0.286	0.317	0.262
GPa	0.099	0.084	0.095	0.093	0.088	0.094	0.075	0.087
GEV	0.063	0.060	0.094	0.067	0.090	0.074	0.071	0.074
LN	0.126	0.102	0.121	0.122	0.108	0.116	0.100	0.086
GAM	0.189	0.186	0.169	0.176	0.175	0.177	0.178	0.154
EV	0.403	0.463	0.402	0.460	0.427	0.393	0.417	0.384
LLG	0.103	0.083	0.105	0.101	0.088	0.089	0.080	0.080
STA	0.059	0.063	0.125	0.060	0.092	0.287	0.087	0.613

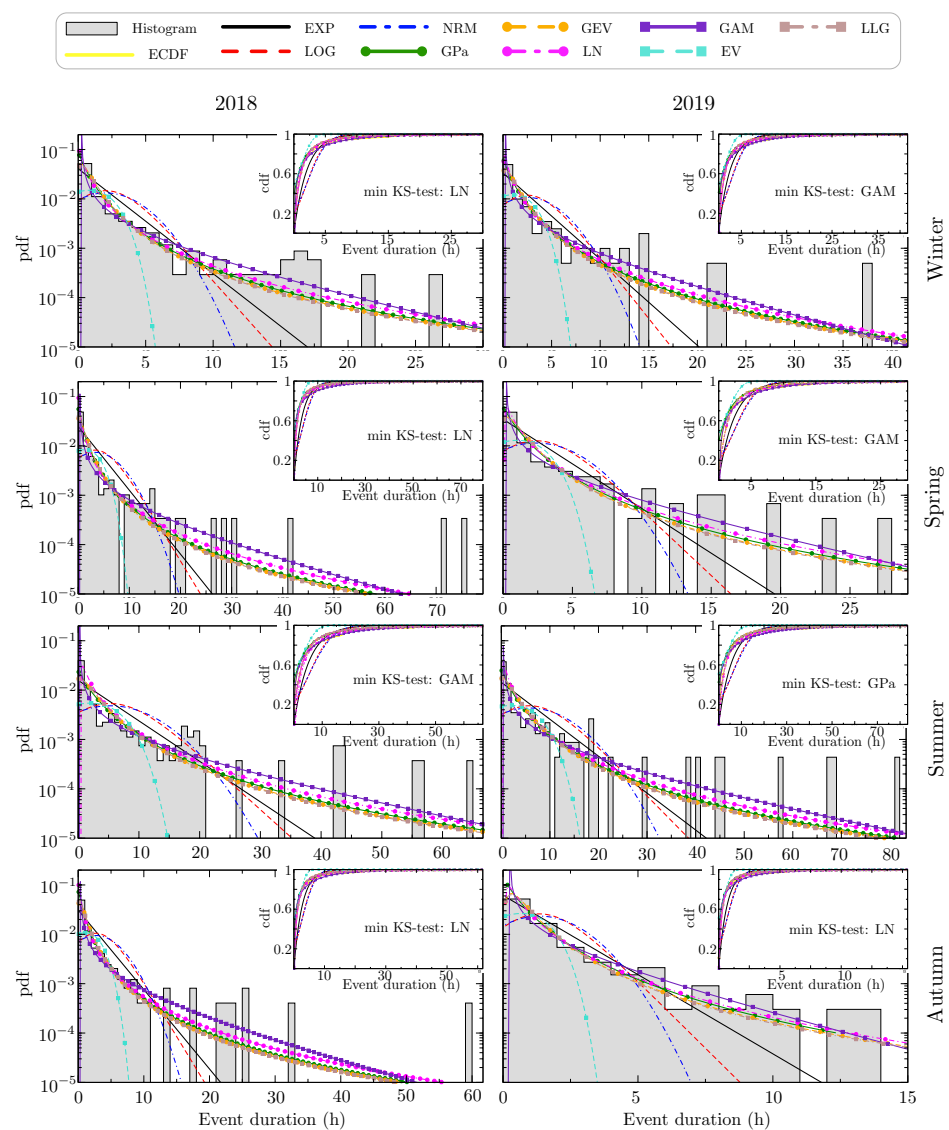




**Figure 5.** Probability density functions of fog events over time for the Maximum Likelihood approach fixing the threshold to 2000 m. The CDF is located in the insets. Each row corresponds to a season and each column to a year.

Table 6 shows the obtained numerical KS-distance using the L-moments method. Furthermore, Figure 6 illustrates the estimated distributions where the y-axis is plotted logarithmically scaled. A total of nine distributions were taken into account in this case, because the moments of the STA distribution do not converge for certain parameters of the distribution. The results show that the distributions that best fit to the fog events in these two years of study are the LN, GAM and GPa with KS-distances around between 0.1 and 0.2. In addition the Log-Logistic (LLG) presents KS-distance stationary values over time, about 0.1. These four distributions are used for modelling hydrological processes or more generally in natural systems [53]. Specifically, the GAM distribution applies to a wide range of physical processes and is related to other distributions: EXP, Pascal, Erlang, Poisson, and chisquare. It is commonly used in meteorological processes i.e., to represent pollutant concentrations and precipitation quantities [54]. Moreover, it is used to measure the time between the occurrence of events when the event process is not completely random [55].

Similarly, in our case, fog events in northwest of Iberian Peninsula, especially in summer, are impulsed by the displacement of the Azores anticyclone. GAM distribution seems to benefit from the L-moments estimation method obtaining lower KS-distances than in case of being estimated by the maximum likelihood method, see Tables 5 and 6. However, from a qualitative point of view, GAM probability density function resembles more of a straight line as we move away towards  $+\infty$  in Figures 5 and 6. This is not the observed behavior of the data distribution. It is expected that, as soon as the number of samples increases, GAM will fit worse to the data distribution. Once again, heavy-tailed distributions are the ones that best fit to these meteorological situations in Mondoñedo, except for GAM which is a light-tailed and also obtains good results. The light-tailed distributions such as EXP, NRM and EV obtain the poorest fitting to the data. See for example Spring 2018 with a KS-distance of 0.517 in EV distribution, and how together with the LOG, EXP and the NRM they do not adjust correctly to the extreme values of two events of more than 70 h located in the tail of the data distribution.



**Figure 6.** Probability density functions of fog events over time for L-moments approach fixing the threshold to 2000 m. The CDF is located in the insets. Each row corresponds to a season and each column to a year.

**Table 6.** KS-test results for the L-moment estimation, and for the 2000 m threshold.

	2018				2019			
	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn
EXP	0.294	0.342	0.340	0.310	0.289	0.282	0.343	0.172
LOG	0.318	0.355	0.327	0.336	0.313	0.320	0.325	0.272
NRM	0.302	0.340	0.313	0.325	0.297	0.307	0.308	0.260
GPa	0.116	0.121	0.169	0.107	0.115	0.119	0.148	0.064
GEV	0.148	0.130	0.161	0.137	0.141	0.152	0.160	0.091
LN	0.097	0.065	0.380	0.075	0.130	0.171	0.248	0.059
GAM	0.144	0.240	0.080	0.223	0.107	0.110	0.215	0.091
EV	0.496	0.517	0.506	0.510	0.498	0.488	0.509	0.449
LLG	0.151	0.131	0.165	0.139	0.145	0.156	0.163	0.097

Observing the obtained results in Table 5, there exist some distributions whose KS-distances hardly vary among seasons, such as, GPa or GEV. This is due to the fact that these distributions explain the low visibility event durations equally well among seasons, even though their durations may change between seasons. On the contrary, those distributions whose KS-distances vary among seasons, such as GAM, cannot adapt to the new conditions by simply changing their parameters.

It is possible to notice some differences between the Maximum Likelihood and L-moments methods in the results previously shown. Note that we obtain slightly better KS-distances in the fittest distributions estimated with the Maximum Likelihood method than with L-moments method. This is the case of GEV, which obtains the best KS results through Maximum Likelihood, below 0.08, see Table 5. However, LN, which best fits the data distribution, obtains KS-distances around 0.1 with a high variance among seasons and years, as can be seen in Table 6. It may seem in some instances that the fitting of the GAM distribution is marginally better than that of the LN in some cases. However, it should be noted that, true to its light-tailed nature, the GAM distribution crosses over all of the heavy-tailed distributions (and specifically that of the LN) at large values of fog duration. Therefore, even though GAM may seem a good fit, it fails at large values of fog duration. This is proof that the Maximum Likelihood method fits the main body of the distribution (thus, failing at large values of fog duration) while the L-moment method fits the tails of the present data (failing at low values of fog duration).

As a final note on this point, an accurate statistical characterization of fogs events with extreme-valued distributions can be used to simulate their occurrence at Mondoñedo, within traffic simulators. This way the real effects of deep fog events on traffic causing jams and important circulation problems can be studied.

### 3.1.1. Discussion: Physical Mechanism

The data used in this study show that, although the number of low visibility events is quite similar in all seasons (see Table 4) they last longer in the warm season. The explanation for this pattern can be found in the high pressure system most influential for the Atlantic and Europe: the Azores Anticyclone. In summer, this pressure system strengthens and reaches its most northerly position [56], bringing northerly winds to the Iberian Peninsula. As have been discussed by some authors [57,58] the main ingredients needed for the formation of the low visibility events that affect the “Alto de O Fiouco” area are northward winds that push warm and humid air masses coming from the sea. Since in this region there is a large mountain barrier of around 600–700 m pretty close to both, the Atlantic Ocean and the Cantabrian Sea, these parcels of air are lifted adiabatically becoming saturated at relatively low levels. In addition, the presence of the typical subsidence inversion caused by the Azores Anticyclone forces the formation of low level layers of clouds (mainly stratus and stratocumulus) which can affect the A-8 motor-road since its elevation at this specific location is quite similar to the level at which these clouds are formed (the so-called lifting condensation level). Furthermore, it should be noted that because this meteorological phenomena is caused by a maritime air mass the number of hygroscopic particles (mainly

sea salt) can be considerably higher than normal, which can play a major role in the formation of dense fog events [58].

### 3.1.2. Statistical Study with Different Thresholds for Defining Low-Visibility Events

The results on statistical characterization of low-visibility fog episodes shown above consider as low-visibility events those under the limit of the visibilimeter (<2000 m, light low-visibility). However, note that this threshold to consider a fog event as low-visibility can be set by the practitioner at alternative values. For example, we can choose different thresholds related to traffic protocols, such as 600 m (moderate low-visibility), 300 m (severe low-visibility), and 50 m (extremely severe low-visibility), all of them with an important effect in secure driving conditions. In fact, visibility below 50 m (extremely severe) very probably leads to motor-road closure. Tables 7 and 8 show the statistics for the low-visibility events at A8 motor-road in the years of the study (2018 and 2019), considering low-visibility those events under 600 m and 300 m, respectively. As can be seen, the number of low-visibility events with the new thresholds is similar, between them and also to the case of the threshold set at 2000 m (light low-visibility events). The low visibility event durations are also quite similar for these three thresholds and their proportions among the seasons remain constant. The low visibility events in the warmer seasons continue to be the longest lasting, see Tables 4 and 7–9. This is due to the fact that fog events are usually very intense at Mondoñedo in this season. Table 9 shows the case of the threshold at 50 m. In this extreme severe case, the number of events is very reduced with respect to other thresholds. This indicates that extreme severe low-visibility events are less frequent than moderate and severe events, mainly in winter and autumn, but with a significant incidence in spring and summer.

**Table 7.** Low-visibility events statistics at Mondoñedo station for the years 2018 and 2019, and for the 600 m threshold.

	Fog Events Average Duration (Minutes)		# of Fog Events	
	2018	2019	2018	2019
Winter	102.17	116.72	278	180
Spring	187.32	135.11	237	224
Summer	274.77	283.10	240	220
Autumn	126.67	61.99	226	260

**Table 8.** Low-visibility events statistics at Mondoñedo station for the years 2018 and 2019, for the 300 m threshold.

	Fog Events Average Duration (Minutes)		# of Fog Events	
	2018	2019	2018	2019
Winter	97.41	98.90	256	186
Spring	194.56	124.62	235	222
Summer	249.76	293.95	254	204
Autumn	138.47	52.63	185	239

**Table 9.** Low-visibility events statistics at Mondoñedo station for the years 2018 and 2019, and for the 50 m threshold.

	Fog Events Average Duration (Minutes)		# of Fog Events	
	2018	2019	2018	2019
Winter	55.43	63.12	34	15
Spring	107.90	60.97	186	133
Summer	156.10	151.48	263	228
Autumn	119.39	16.25	97	3

We repeat here the analysis of the probability distributions for fog events duration, considering low-visibility events defined by setting the thresholds to 600, 300 and 50 m. We consider ten distributions including light and heavy tail distributions, using both the

Maximum Likelihood and L-moments methods. In the case of Maximum Likelihood estimation, Tables 10–12 show the obtained KS-distance for each distribution, divided by seasons and years. We clearly distinguish two different statistic behaviors in the results obtained. Low-visibility events defined by thresholds under 600 and 300 m have a very similar behavior than that defined by a threshold at 2000 m, as can be seen in Tables 10 and 11, respectively. Figure 7 shows the distributions estimated by the Maximum Likelihood method for all seasons in 2018 and 2019 at the 300 m threshold which helps us along the discussion. GEV, GPa and STA are still the distributions which best fit the data, with KS-distances below 0.1 for both the 600 and 300 m thresholds. Their good approximation to the data distribution is clearly presented in Figure 7 for the 300 m threshold. The non-negligible probability of the extreme events is responsible of the good results reported by these heavy-tail distribution, similar to those obtained with a 2000 m threshold, see Table 5 and Figure 5. GEV obtains the best KS-distances in most of seasons of the two years analyzed, closely followed by STA. Both distributions report KS-distances around 0.08 in most seasons. Even for autumn 2019 at the 300 m threshold which is a season with no extreme fog events, see Table 11 and Figure 7. However, STA fails to fit data distribution in both summers, where most of the extremes take place. LLG and GPa obtain larger KS-distances than the previous one between 0.09 and 0.11, but still with good results. The results of the rest of the evaluated distributions are far from these ones discussed previously, especially that provided by the light-tail distributions EXP, LOG, NRM and EV, which are characterized by a quick decrease in probability. In Figure 7, we see that EXP and LOG distributions are straight lines with different slopes, mainly far from the origin, and EV and NRM have concave shape, which does not fit the data distribution trend. The KS-distances obtained by these distributions are above 0.3 for both thresholds. In the case of low-visibility events defined by the threshold of 50 m, the behavior changes slightly with respect to the case of the threshold situated at 2000 m, see Table 12. Again, the best KS-distances are obtained by the heavy-tail distributions GEV, and STA, but their KS distances are now around 0.1 with more variations among the seasons. The light-tail distributions still obtain the worst KS-distances, but they decrease respect to the previous threshold. We cannot estimate distributions of autumn 2019 for a threshold of 50 m since the number of low-visibility events in this season is only 3.

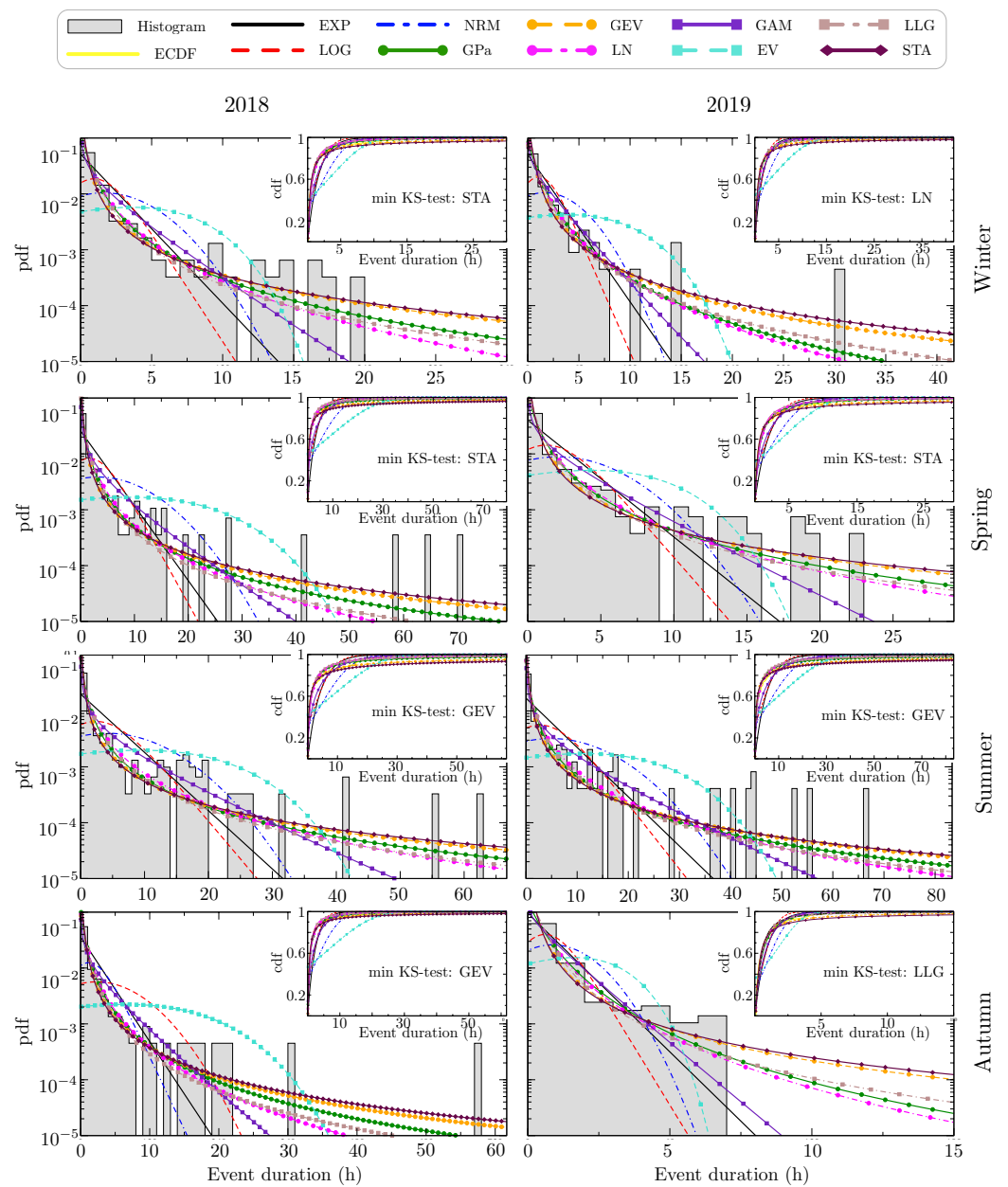
**Table 10.** KS-distances results for the Maximum likelihood, for the 600 m threshold.

	2018				2019			
	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn
EXP	0.279	0.383	0.381	0.334	0.251	0.262	0.388	0.195
LOG	0.313	0.345	0.327	0.338	0.293	0.298	0.333	0.269
NRM	0.310	0.355	0.318	0.355	0.313	0.295	0.326	0.271
GPa	0.117	0.113	0.088	0.120	0.097	0.089	0.088	0.123
GEV	0.073	0.072	0.057	0.091	0.105	0.073	0.081	0.083
LN	0.101	0.111	0.110	0.088	0.111	0.074	0.097	0.106
GAM	0.171	0.202	0.187	0.190	0.137	0.147	0.193	0.153
EV	0.401	0.465	0.415	0.476	0.448	0.400	0.417	0.374
LLG	0.088	0.093	0.085	0.089	0.106	0.071	0.083	0.093
STA	0.071	0.068	0.436	0.086	0.106	0.072	0.495	0.087

**Table 11.** KS-distances results for the Maximum likelihood, for the 300 m threshold.

	2018				2019			
	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn
EXP	0.279	0.395	0.386	0.333	0.205	0.287	0.379	0.198
LOG	0.317	0.351	0.327	0.339	0.276	0.307	0.329	0.267
NRM	0.313	0.357	0.312	0.352	0.310	0.297	0.320	0.267
GPa	0.125	0.118	0.103	0.110	0.092	0.107	0.090	0.138
GEV	0.089	0.080	0.081	0.067	0.073	0.076	0.066	0.105
LN	0.115	0.124	0.118	0.107	0.070	0.098	0.100	0.104
GAM	0.158	0.200	0.186	0.196	0.139	0.166	0.175	0.159
EV	0.403	0.457	0.422	0.469	0.450	0.389	0.411	0.366
LLG	0.104	0.100	0.100	0.092	0.072	0.087	0.088	0.096
STA	0.083	0.078	0.311	0.070	0.078	0.072	0.181	0.101





**Figure 7.** Probability density functions of fog events over time for the Maximum Likelihood approach fixing the threshold to 300 m. The CDF is located in the insets. Each row corresponds to a season and each column to a year.

**Table 12.** KS-distances results for the Maximum likelihood approach, for the 50 m threshold. Estimations of autumn 2019 are not possible since there are only three low-visibility events with a threshold under 50 m in that period.

	2018				2019			
	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn
EXP	0.232	0.322	0.347	0.305	0.355	0.243	0.332	--
LOG	0.258	0.329	0.326	0.308	0.305	0.282	0.318	--
NRM	0.299	0.342	0.317	0.303	0.358	0.262	0.301	--
GPa	0.129	0.135	0.117	0.102	0.243	0.144	0.117	--
GEV	0.131	0.124	0.095	0.067	0.192	0.132	0.110	--
LN	0.145	0.108	0.117	0.105	0.256	0.123	0.117	--
GAM	0.193	0.177	0.177	0.190	0.316	0.176	0.172	--
EV	0.409	0.462	0.420	0.390	0.356	0.342	0.388	--
LLG	0.138	0.106	0.102	0.088	0.227	0.111	0.106	--
STA	0.131	0.118	0.090	0.063	0.181	0.125	0.101	--

Tables 13–15 present the KS-distances obtained in cases when the L-moments method is used for estimating the distributions, with thresholds at 600, 300 and 50 m, respectively. Figure 8 shows the distributions estimated by the L-moments method for all seasons in both 2018 and 2019 at the 300 m threshold. The distributions that best fits to the fog events for the 600 and 300 m thresholds is still the LN, and GPa but with higher values respect to the 2000 m threshold, around 0.1, and 0.13, respectively, see Tables 13 and 14. GAM obtains good KS-distances at the 300 m threshold but not in 600 m and varies along the seasons. Furthermore, GAM struggles to fit data in spring and autumn 2018 due to the used L-moments implementation code. The reason is that the duration of low-visibility events in summer is higher than in other seasons, and GAM did fit such wide range of durations with a good accuracy, since they are straight lines far from the origin in the y-log-scaled Figure 8. GEV also fits were to the data distribution even better than fixing a 2000 m threshold, around 0.13. The light-tail distributions do not fit the data; although, they obtain better KS-distance than in the case of the 2000 m threshold. EXP, LOG, NRM or EV do not fit the data distribution well as their tail decreases quickly, see Figure 8. Focusing on the results obtained by fixing a threshold at 50 m, Table 12 shows similar results to the previous thresholds for the L-moment estimation. Again, the distributions with best KS-distances are GAM, LN and GPA, with values above 0.11, higher than in previous thresholds. However, the KS-distances obtained by EXP, EV, NRM are lower than the obtained by fixing 600 and 300 m thresholds. Note that distribution estimations for autumn 2019 do not appear in Table 12, since only three extreme low-visibility events occurred, not enough for the parameter estimation.

**Table 13.** KS-test results for the L-moment estimation, for the 600 m threshold.

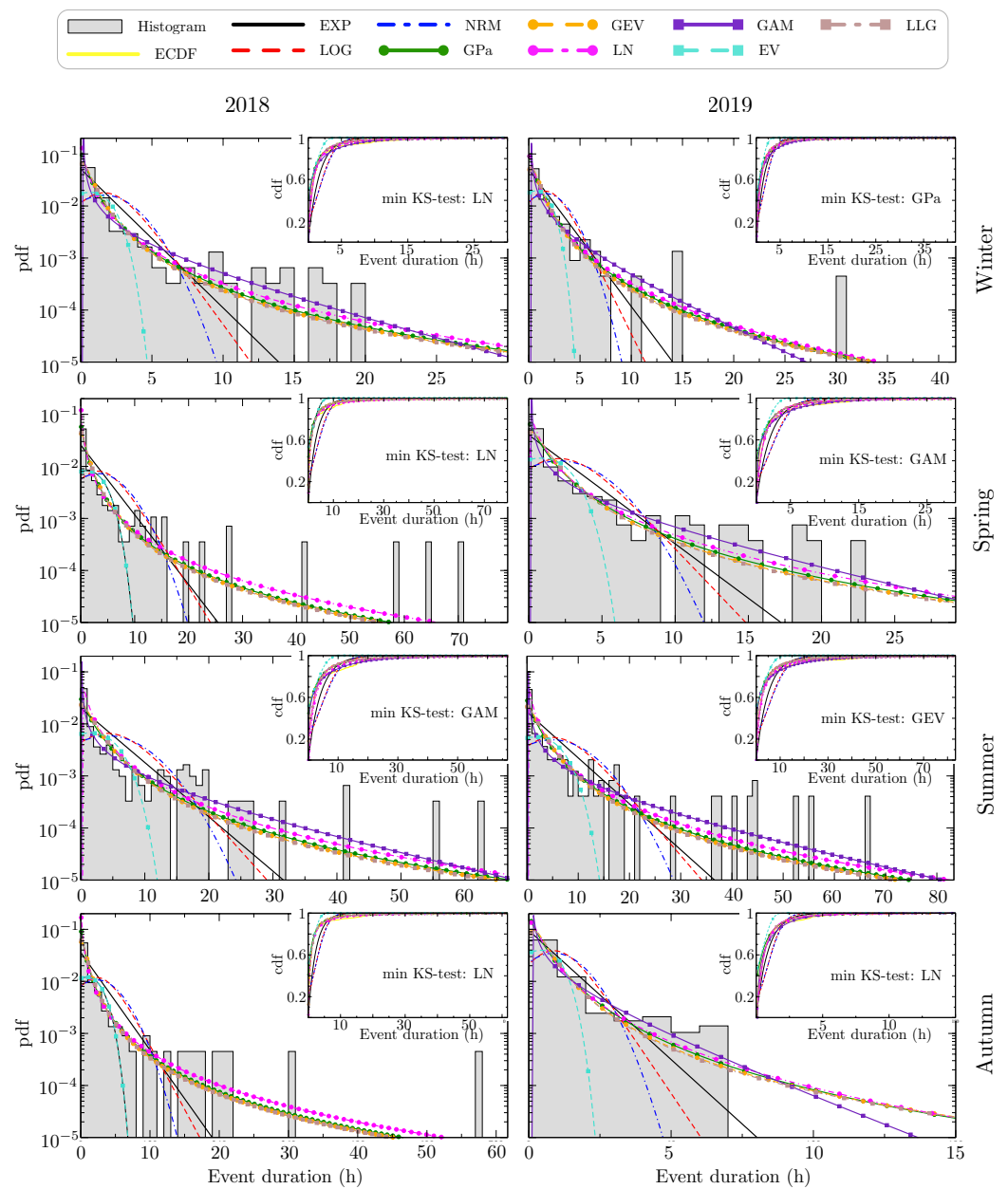
	2018				2019			
	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn
EXP	0.279	0.383	0.381	0.334	0.251	0.262	0.388	0.195
LOG	0.329	0.372	0.349	0.356	0.299	0.309	0.353	0.288
NRM	0.319	0.358	0.334	0.343	0.291	0.298	0.336	0.273
GPa	0.124	0.166	0.184	0.129	0.122	0.128	0.182	0.087
GEV	0.129	0.163	0.184	0.137	0.125	0.129	0.174	0.122
LN	0.094	0.099	0.316	0.069	0.161	0.162	0.279	0.082
GAM	0.241	--	0.179	--	0.261	0.192	0.214	0.092
GEV	0.504	0.529	0.518	0.522	0.495	0.495	0.523	0.471
LLG	0.132	0.164	0.186	0.138	0.129	0.133	0.176	0.127

**Table 14.** KS-test results for the L-moment estimation, for the 300 m threshold.

	2018				2019			
	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn
EXP	0.279	0.395	0.386	0.333	0.205	0.287	0.379	0.198
LOG	0.334	0.372	0.351	0.350	0.290	0.326	0.333	0.265
NRM	0.323	0.360	0.337	0.341	0.277	0.315	0.318	0.253
GPa	0.133	0.174	0.197	0.110	0.077	0.148	0.189	0.106
GEV	0.135	0.170	0.193	0.125	0.098	0.144	0.176	0.123
LN	0.104	0.108	0.411	0.052	0.057	0.193	0.341	0.093
GAM	0.121	--	0.106	--	0.199	0.104	0.211	0.146
EV	0.507	0.533	0.518	0.522	0.481	0.502	0.521	0.472
LLG	0.138	0.171	0.195	0.125	0.102	0.147	0.178	0.130

**Table 15.** KS-test results for the L-moment estimation, for the 50 m threshold.

	2018				2019			
	Winter	Spring	Summer	Autumn	Winter	Spring	Summer	Autumn
EXP	0.232	0.322	0.347	0.305	0.355	0.243	0.332	--
LOG	0.267	0.359	0.337	0.327	0.398	0.299	0.328	--
NRM	0.250	0.349	0.322	0.313	0.381	0.285	0.316	--
GPa	0.113	0.153	0.177	0.120	0.242	0.156	0.183	--
GEV	0.144	0.151	0.166	0.134	0.274	0.146	0.169	--
LN	0.114	0.121	0.238	0.094	0.334	0.261	0.320	--
GAM	0.088	0.161	0.125	0.175	0.176	0.173	0.140	--
EV	0.477	0.520	0.516	0.502	0.472	0.484	0.511	--
LLG	0.160	0.153	0.169	0.142	0.304	0.154	0.172	--



**Figure 8.** Probability density functions of fog events over time for the L-moments approach establishing the threshold at 300 m. The CDF is located in the insets. Each row corresponds to a season and each column to a year.

### 3.2. Prediction of Fog Events at A-8 with ELMs

The results obtained with an ELM in the short-term prediction of low-visibility events due to fog at the A-8 motor-road are presented in this section. For ensuring the independence of the partition data in training and test sets, as well as the performance of the regressors, a *K*-fold cross-validation procedure was carried out [10,59]. The folding was set to *K* = 10, and each set consists of an 80% to train and 20% to test. Using the full dataset spanning from 1 January 2018 to 30 November 2019, data are randomly selected, breaking the sequence in the data, in order to bring heterogeneity to the values of the samples.

The ELM model considered in this paper has the following characteristics: neurons in the hidden layer are designed with sigmoid activation function. The optimal number of neurons is chosen from a large pool (50–150, in an increment of 1), which passes through the hidden layer, one by one, during the validation phase. In addition to the atmospheric features considered, we will also use the 4 time instants prior to the target we want to

predict ( $t$ ), as predictors, i.e., we will use the target values at  $t - 1$ ,  $t - 2$ ,  $t - 3$  and  $t - 4$ . We should note that in all experiments the input–output data pair,  $\{(x_i, y_i)\}_{i=1}^n$ , for both ELMs has a time resolution of half an hour (where  $n$  stands for the total number of half-hour intervals in the database); hence, the forecasting time-horizon was set to 30 min ahead estimation (instant  $t$ ) of the visibility. Finally, the experiments will consist of launching 10 executions of each algorithm for each proposed scenario, and average the results of them.

In order to better analyze the ELM performance, a wrapper feature selection process was carried out. This procedure consists of launching as many ELMs as combinations of characteristics we have in a reduced validation set, to find the set of predictors that provides the least error at the output (best set of features). Note that we have 10 features (only the atmospheric features are considered in this process) for this problem (see Table 1), which means that we have to launch a total of 1024 ( $2^{10}$ ) ELM models (prediction problems) to obtain the best set of characteristics (inputs). Note that we need an extremely fast-training algorithm such as ELM to carry out this feature selection analysis, since otherwise the computation time required would be extremely high. The results obtained in the feature selection process provided two sets of features as best results: the first one included a total of nine characteristics: Accumulated precipitation, Salinity, Relative humidity, Air temperature, Floor temperature, Dew temperature, Global solar radiation, Wind speed and Atmospheric pressure, with a Root Mean Square error (RMSE) of 378.54 m in the validation set. A second best set with a total of three characteristics was also obtained: Accumulated precipitation, Relative humidity and Global solar radiation; with a RMSE of 381.36 m in the validation set. The rest of features combination produced worse results, so we have kept these two best sets for carrying out the experiments. In order to compare the results, we used the Persistence Prediction Operator ( $PPO_1$ ), a well known operator described by the following equation:

$$x(t + 1) = x(t). \quad (6)$$

The generalized Persistence prediction operator ( $PPO_M$ ) uses the  $M$  last time steps to infer the prediction and can be also defined as:

$$x(t + 1) = \frac{1}{M} \sum_{i=1}^M x(t - i). \quad (7)$$

In our experiments, we fix  $M = 4$  referenced as ( $PPO_4$ ).

### Prediction Results

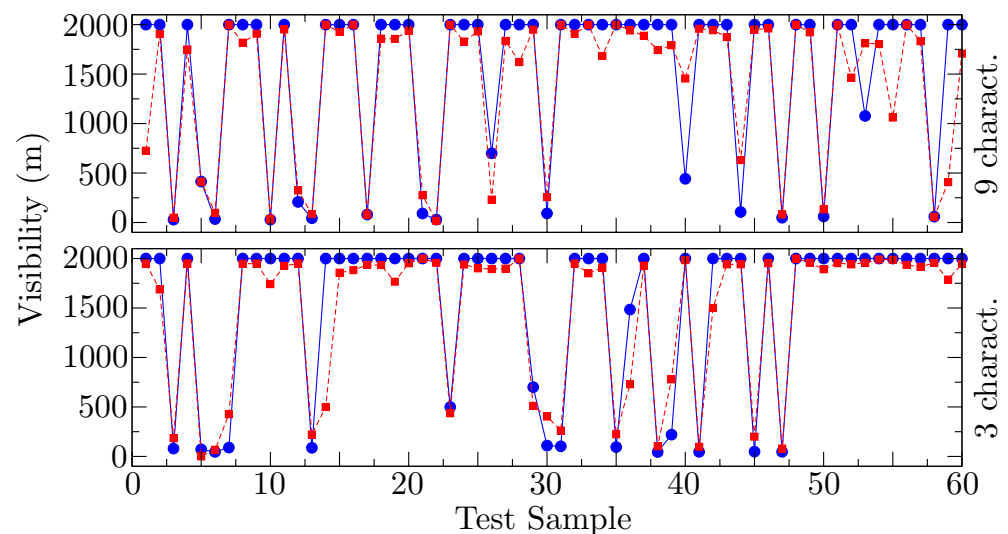
Table 16 shows the average and standard deviation results (10 runs of the algorithms) obtained by the ELM, when we use 9 or 3 characteristics as predictors (ELM-9, ELM-3), and the obtained by the PPO, which uses the last ( $PPO_1$ ) and the four last time steps ( $PPO_4$ ) for comparison. These results to evaluate the ELM performance are given in terms of the Pearson's correlation coefficient,  $r^2$ , the RMSE and the computation times, both training (Train- $t$ ) and test (Test- $t$ ). It can be observed that the ELM approach is able to obtain the best results, with an RMSE of 393.56 m and an 80% of  $r^2$  when 3 features are used as predictors. If we compare these results with the case of ELM-9, we can observe a slight difference in terms of RMSE, with a value of 394.82 m, but a similar value for  $r^2$ . Therefore, the ELM model works slightly better with fewer features, achieving good results using less computation time, in particular taking 16.06 and 0.03 s in Train- $t$  and Test- $t$ , respectively in ELM-3, against the 18.99 and 0.05 s for Train- $t$  and Test- $t$  in the ELM-9 case. Based on these results we can see that the selection of features has an effect on the prediction process by using the ELM approach. The results obtained by the PPO are considerably worse. If we analyze the PPO for both variants,  $PPO_1$  and  $PPO_4$ , we can see the poor performance in terms of RMSE and the worse one in  $r^2$ . In this case, the best results are obtained for  $PPO_1$  with an RMSE of 418.09 m and 0.77 in  $r^2$ , below the 0.80 of the ELM. The difference with respect to the case of  $PPO_4$  is larger than in the ELM at least in terms of RMSE which obtains 477.65 m. In terms of  $r^2$ ,  $PPO_4$  obtains 0.72, similar to the  $PPO_1$ . We deduce that,

in case of PPO, it is better to use the last time step than four time steps for obtaining the prediction. This is because the visibility time series is quite volatile, and using  $PPO_4$  (7) strongly smooths the time series. Note that the computation time required to train the ELM is acceptable. PPO reaches real time as the predicted series is simply the mean given by Equation (7).

Figure 9 shows a temporal representation of the predicted visibility variable (in red) versus the measured values of this variable (in blue), by the ELM with nine features and three features. It is possible to see that in both cases the performance of the ELM is excellent in this prediction problem, showing good behaviour even in the deepest fog events. Moreover, these good results are obtained regardless of the set of samples we test. As can be seen, in the prediction graphs with nine and three features as predictors, different test samples are used to quantify the performance of the model, which is very interesting to corroborate the good performance of this type of learning machine.

**Table 16.** Prediction error results by the ELM and PPO approaches, and computational running time (Train-*t* and Test-*t*) by ELM.

	ELM-9		ELM-3		PPO <sub>1</sub>		PPO <sub>4</sub>	
	Avg	Std	Avg	Std	Avg	Std	Avg	Std
$r^2$	0.80	0.01	0.80	0.01	0.77	–	0.72	–
RMSE (meters)	394.82	9.50	393.56	5.28	418.09	–	477.65	–
Train- <i>t</i> (seconds)	18.99	1.00	16.06	0.65	–	–	–	–
Test- <i>t</i> (seconds)	0.05	0.01	0.03	0.004	–	–	–	–



**Figure 9.** Time-series of measured (blue) and predicted (red) visibility values by the ELM with 9 and 3 features (input variables).

#### 4. Conclusions

In this paper we carried out a detailed study on low-visibility events due to fog at the A-8 motor-road, Galicia, Spain. In this zone there are frequent episodes of low-visibility events due to orographic fog from the Cantabrian sea, which deeply affect the traffic in the road. First, we statistically characterized the fog events’ duration with different distributions, including light and heavy-tail distributions, in a seasonal analysis. Two different approaches were considered to estimate the parameters associated with the distributions, the Maximum Likelihood criterion and the L-moments approach. In both cases we showed that the heavy-tail distributions (extreme distributions) are more suitable to describe the duration of low-visibility events at the A-8 motor-road, independently of the season considered. We can state that three of the heavy-tailed distributions are the best fit in most cases. These are the GEV, the GP<sub>a</sub> and LN. These three distributions present the



smallest KS-distances values in most cases, especially in those where this phenomenon of dense and long-lasting fogs takes place: in summer. LLG also obtains consistently very good results closed to the previous ones among all seasons of the studied years for both estimation methods. Other distributions such as GAM or STA offer good results in some cases, but this is not the majority. In some cases, GAM reports very good KS-distances when is estimated by the L-moment method, not as good by the maximum likelihood method, even as a light-tailed distribution. We carried out an analysis by considering different visibility thresholds to define the low-visibility events, showing that the results are very similar in light, moderate and severe low-visibility events situations (thresholds of 2000 m, 600 m and 300 m), whereas for extreme severe low-visibility events (less than 50 m), their statistical characterization is different.

Second, we tackled the prediction of the visibility in the zone, by using an ELM regressor. Atmospheric variables collected in-situ such as air and floor temperature, wind speed and direction, pressure, etc. were used as inputs to carry out the prediction. We showed an accurate performance of the ELM with average errors under 400 m in all cases, improving by far the performance of the persistence prediction operator in the problem. In all the experiments carried out, the variance of the prediction results as very low, which indicates that the prediction is accurate, i.e., the ELM obtains a good prediction, even for different low-visibility ranges (light, moderate, severe and extreme severe low-visibility events).

The methods and algorithms presented in this paper have a direct application in dealing with the issues caused by low-visibility events from orographic fog on the A-8 motor-road. The statistical description of the events allows an accurate modeling of low-visibility event's duration in traffic models simulations. This can be complimented with and accurate prediction of low-visibility events, in order to implement actions to minimize their impact in the traffic of the zone.

Note that some limitations are found in this work, mainly in the capacity of the proposed methods to characterize and predict extreme low-visibility events. Future analysis of low-visibility events in the zone can include numerical weather prediction approaches combined with ML algorithms, to help deal with these cases. Numerical methods such as meso-scale models can improve the accuracy of the prediction models, by including new information as input variables. In turn, we can complete real data with numerical weather models simulations, in order to produce larger datasets to improve the accuracy of the statistical models for low-visibility events' description.

**Author Contributions:** Conceptualization, S.S.-S., C.C.-M., A.J.C.; methodology, S.S.-S. and D.C.-P.; software, S.C.-B., L.C.-B., M.I.C. and D.C.-P.; validation, S.C.-B., L.C.-B., M.I.C., E.C.-P.; data curation, E.C.-P. and C.C.-M.; writing—original draft preparation, S.C.-B., S.S.-S., D.C.-P., A.J.C.; writing—revised draft, S.S.-S., D.C.-P., S.C.-B.; supervision, S.S.-S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** This research has been partially supported by the Ministerio de Economía, Industria y Competitividad of Spain (Grant Ref TIN2017-85887-C2-2-P and TIN2017-90567-REDT). The authors would like to thank GSJ SOLUTIONS, part of SANJOSE Group, for providing the A8 weather data.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Fabbian, D.; De-Dear, R.; Lellyett, S. Application of Artificial Neural Network Forecasts to Predict Fog at Canberra International Airport. *Weather Forecast.* **2007**, *22*, 372–381. [[CrossRef](#)]

2. Miao, Y.; Potts, R.; Huang, X.; Elliott, G.; Rivett, R. A fuzzy logic fog forecasting model for Perth airport. *Pure Appl. Geophys.* **2012**, *169*, 1107–1119. [[CrossRef](#)]
3. Guerreiro, P.M.; Soares, P.M.; Cardoso, R.M.; Ramos, A.M. An Analysis of Fog in the Mainland Portuguese International Airports. *Atmosphere* **2020**, *11*, 1239. [[CrossRef](#)]
4. Peng, Y.; Abdel-Aty, M.; Lee, J.; Zou, Y. Analysis of the impact of fog-related reduced visibility on traffic parameters. *J. Transp. Eng. Part A Syst.* **2018**, *144*, 04017077. [[CrossRef](#)]
5. Wu, Y.; Abdel-Aty, M.; Lee, J. Crash risk analysis during fog conditions using real-time traffic data. *Accid. Anal. Prev.* **2018**, *114*, 4–11. [[CrossRef](#)] [[PubMed](#)]
6. Musk, L.; Perry, A. Climate as a factor in the planning and design of new roads and motorways. *Highw. Meteorol.* **1991**, *59*, 1–25.
7. Cho, H.J.; Kim, K.S. Development of hazardous road fog index and its application. *J. East. Asia Soc. Transp. Stud.* **2005**, *6*, 3357–3371.
8. Peón, D.; Rodríguez-Álvarez, J.; López-Iglesias, E. Spread or backwash: The impact on population dynamics and business performance of a new road in a rural county of Galicia (Spain). *Pap. Reg. Sci.* **2019**, *98*, 2479–2502. [[CrossRef](#)]
9. Räsänen, M.; Chung, M.; Katurji, M.; Pellikka, P. Similarity in Fog and Rainfall Intermittency. *Geophys. Res. Lett.* **2018**, *45*, 10691–10699. [[CrossRef](#)]
10. Cornejo-Bueno, S.; Casillas-Pérez, D.; Cornejo-Bueno, L.; Chidean, M.I.; Caamaño, A.J.; Sanz-Justo, J.; Casanova-Mateo, C.; Salcedo-Sanz, S. Persistence Analysis and Prediction of Low-Visibility Events at Valladolid Airport, Spain. *Symmetry* **2020**, *12*, 1045. [[CrossRef](#)]
11. Salcedo-Sanz, S.; Piles, M.; Cuadra, L.; Casanova-Mateo, C.; Caamaño, A.; Cerro-Prada, E.; Camps-Valls, G. Long-term persistence, invariant time scales and on-off intermittency of fog events. *Atmos. Res.* **2021**, *252*, 105456. [[CrossRef](#)]
12. Price, J.; Stokkerei, K. The Use of Thermal Infra-Red Imagery to Elucidate the Dynamics and Processes Occurring in Fog. *Atmosphere* **2020**, *11*, 240. [[CrossRef](#)]
13. van der Velde, I.R.; Steeneveld, G.J.; Wichers Schreur, B.G.J.; Holtslag, A.A.M. Modeling and Forecasting the onset and duration of severe radiation fog under frost conditions. *Mon. Weather Rev.* **2010**, *138*, 4237–4253. [[CrossRef](#)]
14. Román-Cascón, C.; Yagüe, C.; Sastre, M.; Maqueda, G.; Salamanca, F.; Viana, S. Observations and WRF simulations of fog events at the Spanish Northern Plateau. *Adv. Sci. Res.* **2012**, *8*, 11–18. [[CrossRef](#)]
15. Akimoto, Y.; Kusaka, H. A climatological study of fog in Japan based on event data. *Atmos. Res.* **2015**, *151*, 200–211. [[CrossRef](#)]
16. Stolaki, S.; Haeffelin, M.; Lac, C.; Dupont, J.C.; Elias, T.; Masson, V. Influence of aerosols on the life cycle of a radiation fog event. A numerical and observational study. *Atmos. Res.* **2015**, *151*, 146–161. [[CrossRef](#)]
17. M. Belo-Pereira, J.A.S. A persistent wintertime fog episode at Lisbon airport (Portugal): Performance of ECMWF and AROME models. *Meteorol. Appl.* **2016**, *23*, 353–370. [[CrossRef](#)]
18. La, I.; Yum, S.S.; Gulpepe, I.; Yeom, J.M.; Song, J.I.; Cha, J.W. Influence of Quasi-Periodic Oscillation of Atmospheric Variables on Radiation Fog over A Mountainous Region of Korea. *Atmosphere* **2020**, *11*, 230. [[CrossRef](#)]
19. Bergot, T.; Terradellas, E.; Cuxart, J.; Mira, A.; Liechti, O.; Mueller, M.; Nielsen, N.W. Intercomparison of Single-Column Numerical Models for the Prediction of Radiation Fog. *J. Appl. Meteorol. Climatol.* **2007**, *46*, 504–521. [[CrossRef](#)]
20. Gulpepe, I.; Tardif, R.; Michaelides, S.C.; Cermak, J.; Bott, A.; Bendix, J.; Müller, M.D.; Pagowski, M.; Hansen, B.; Ellrod, G.; et al. Fog research: A review of past achievements and future perspectives. *Pure Appl. Geophys.* **2007**, *164*, 1121–1159. [[CrossRef](#)]
21. Fernández-González, S.; Bolgiani, P.; Fernández-Villares, J.; González, P.; García-Gil, A.; Suárez, J.C.; Merino, A. Forecasting of poor visibility episodes in the vicinity of Tenerife Norte Airport. *Atmos. Res.* **2019**, *223*, 49–59. [[CrossRef](#)]
22. Smith, D.K.; Renfrew, I.A.; Dorling, S.R.; Price, J.D.; Boutle, I.A. Sub-km scale numerical weather prediction model simulations of radiation fog. *Q. J. R. Meteorol. Soc.* **2020**, *147*, 746–763. [[CrossRef](#)]
23. Zhou, B.; Du, J.; Gulpepe, I.; Dimego, G. Forecast of low visibility and fog from NCEP: Current status and efforts. *Pure Appl. Geophys.* **2011**, *169*, 895–909. [[CrossRef](#)]
24. Koziara, M.; Robert, J.; Thompson, W. Estimating Marine Fog Probability Using a Model Output Statistics Scheme. *Mon. Weather Rev.* **1983**, *111*, 2333–2340. [[CrossRef](#)]
25. Colabone, R.O.; Ferrari, A.; da Silva-Vecchia, F.; Bruno-Tech, A. Application of artificial neural networks for fog forecast. *J. Aerosp. Technol. Manag.* **2015**, *169*, 1107–1119. [[CrossRef](#)]
26. Boneh, T.; Weymouth, G.; Newham, P.; Potts, R.; Bally, J.; Nicholson, A.; Korb, K. Fog Forecasting for Melbourne Airport Using a Bayesian Decision Network. *Weather Forecast.* **2015**, *30*, 1218–1233. [[CrossRef](#)]
27. Cornejo-Bueno, L.; Casanova-Mateo, C.; Sanz-Justo, J.; Cerro-Prada, E.; Salcedo-Sanz, S. Efficient prediction of low-visibility events at airports using Machine-Learning regression. *Bound. Layer Meteorol.* **2017**, *165*, 349–370. [[CrossRef](#)]
28. Zhu, X.; Ni, Z.; Cheng, M.; Jin, F.; Li, J.; Weckman, G. Selective ensemble based on extreme learning machine and improved discrete artificial fish swarm algorithm for haze forecast. *Appl. Intell.* **2018**, *48*, 1757–1775. [[CrossRef](#)]
29. Bari, D. Visibility prediction based on kilometeric nwp model outputs using machine-learning regression. In Proceedings of the 2018 IEEE 14th International Conference on e-Science (e-Science), Amsterdam, The Netherlands, 29 October–1 November 2018; p. 278.
30. Bari, D.; Ouagabi, A. Machine-learning regression applied to diagnose horizontal visibility from mesoscale NWP model forecasts. *SN Appl. Sci.* **2020**, *2*, 1–13. [[CrossRef](#)]

31. Pan, H.; Xue, J.; Huang, M.; Lei, X. Air Visibility Prediction Based on Multiple Models. In Proceedings of the 2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), Tianjin, China, 19–23 July 2018; pp. 1421–1426.
32. Durán-Rosal, A.; Fernández, J.; Casanova-Mateo, C.; Sanz-Justo, J.; Salcedo-Sanz, S.; Hervás-Martínez, C. Efficient fog prediction with multi-objective evolutionary neural networks. *Appl. Soft Comput.* **2018**, *70*, 347–358. [[CrossRef](#)]
33. Guijo-Rubio, D.; Gutiérrez, P.; Casanova-Mateo, C.; Sanz-Justo, J.; Salcedo-Sanz, S.; Hervás-Martínez, C. Prediction of low-visibility events due to fog using ordinal classification. *Atmos. Res.* **2018**, *214*, 64–73. [[CrossRef](#)]
34. Castillo, E. *Extreme Value Theory in Engineering*; Elsevier: Amsterdam, The Netherlands, 2012.
35. Pisarenko, V.; Rodkin, M. The estimation of probability of extreme events for small samples. *Pure Appl. Geophys.* **2017**, *174*, 1547–1560. [[CrossRef](#)]
36. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [[CrossRef](#)]
37. Nolan, J.P. Financial modeling with heavy-tailed stable distributions. *Wiley Interdiscip. Rev. Comput. Stat.* **2014**, *6*, 45–55. [[CrossRef](#)]
38. Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*; US Government Publishing Office: Washington, DC, USA, 1964; Volume 55.
39. Gumbel, E.J. *Statistics of Extremes*; Columbia University Press: New York, NY, USA, 1958; Volume 201.
40. Fisher, R.A. Theory of statistical estimation. In *Mathematical Proceedings of the Cambridge Philosophical Society*; Cambridge University Press: Cambridge, UK, 1925; Volume 22, pp. 700–725.
41. Le Cam, L. Maximum likelihood: An introduction. *Int. Stat. Rev. Int. Stat.* **1990**, *58*, 153–171. [[CrossRef](#)]
42. Hosking, J.R. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. R. Stat. Soc. Ser. B Methodol.* **1990**, *52*, 105–124. [[CrossRef](#)]
43. Nolan, J.P. Numerical calculation of stable densities and distribution functions. *Commun. Stat. Stoch. Model.* **1997**, *13*, 759–774. [[CrossRef](#)]
44. Arnold, B.C.; Balakrishnan, N.; Nagaraja, H.N. *A First Course in Order Statistics*; SIAM: Philadelphia, PA, USA, 2008.
45. Massey, F.J., Jr. The Kolmogorov-Smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **1951**, *46*, 68–78. [[CrossRef](#)]
46. Haykin, S.; Network, N. A comprehensive foundation. *Neural Netw.* **2004**, *2*, 41.
47. Hagan, M.T.; Menhaj, M.B. Training feedforward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.* **1994**, *5*, 989–993. [[CrossRef](#)]
48. Huang, G.B.; Zhou, H.; Ding, X.; Zhang, R. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2011**, *42*, 513–529. [[CrossRef](#)]
49. Autores, V. ELM Matlab Code. Available online: [http://www.ntu.edu.sg/home/egbhuang/elm\\_codes.html](http://www.ntu.edu.sg/home/egbhuang/elm_codes.html) (accessed on 23 May 2021).
50. Teena, N.; Kumar, V.S.; Sudheesh, K.; Sajeev, R. Statistical analysis on extreme wave height. *Nat. Hazards* **2012**, *64*, 223–236. [[CrossRef](#)]
51. Aguilar, R.R.; Ake, S.C. Valuación de opciones de tipo de cambio asumiendo distribuciones  $\alpha$ -estables. *Contaduría Y Adm.* **2013**, *58*, 149–172. [[CrossRef](#)]
52. Campos-Aranda, D.F. Ajuste de las distribuciones GVE, LOG y PAG con momentos L depurados (1, 0). *Tecnol. Y Cienc. Agua* **2015**, *6*, 153–167.
53. El Adlouni, S.; Bobée, B.; Ouarda, T.B. On the tails of extreme event distributions in hydrology. *J. Hydrol.* **2008**, *355*, 16–33. [[CrossRef](#)]
54. Watterson, I.G.; Dix, M.R. Simulated changes due to global warming in daily precipitation means and extremes and their interpretation using the gamma distribution. *J. Geophys. Res. Atmos.* **2003**, *108*. [[CrossRef](#)]
55. Mun, J. *Advanced Analytical Models: Over 800 Models and 300 Applications from the Basel II accord to Wall Street and Beyond*; John Wiley & Sons: Hoboken, NJ, USA, 2008; Volume 419.
56. Davis, R.E.; Hayden, B.P.; Gay, D.A.; Phillips, W.L.; Jones, G.V. The north atlantic subtropical anticyclone. *J. Clim.* **1997**, *10*, 728–744. [[CrossRef](#)]
57. Royé, D.; Rasilla, D.; Marti, A.; Lorenzo, N.; Abalde, N. Análisis espacio-temporal de la nubosidad en el norte de la provincia de Lugo. In Proceedings of the XI Congreso Internacional de la Asociación Española de Climatología (AEC), Cartagena, Spain, 17–19 October 2018.
58. Meteoclim. Consulta Preliminar del Mercado, para la Búsqueda de Soluciones Innovadoras en Proyectos de Innovación Relacionados con el Diseño e Implementación de Sistemas de Protección Antiniebla en la Autovía A-8 entre Mondoñedo y A Xesta, Provincia de Lugo. 2015. Technical Report. Available online: <https://www.mitma.gob.es/> (accessed on 23 May 2021).
59. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai* **1995**, *14*, 1137–1145.