

# Statistical models for the lidar technology: false alarms, receiver operating characteristic curves, and Swerling models

Robert Bernier<sup>1b</sup>,<sup>a,\*</sup> Xiaoying Cao<sup>1b</sup>,<sup>b</sup> Gilles Roy,<sup>c</sup> and Grégoire Tremblay<sup>1b</sup><sup>d</sup>

<sup>a</sup>Les instruments optiques du Saint-Laurent Inc., Mirabel, Quebec, Canada

<sup>b</sup>Lidar Consultant, Ottawa, Ontario, Canada

<sup>c</sup>Lidar Consultant, Shannon, Quebec, Canada

<sup>d</sup>Defence Research Development Canada—Valcartier, Quebec, Canada

**Abstract.** We argue for the integration of the statistical models already widely used in radar technology into lidar technology. The aim is to assess the validity or degree of confidence of an alert to be issued in view of not overloading the pilot with nuisance alerts. We present the basics of the detection theory. We give three examples of simulations illustrating the use of these statistical models either for designing lidars or for preparing lidar missions. We describe the simulator having been developed and used. We also present the idea of developing mixtures of statistical models as an approach to thresholding and object classification at mission time. Some experimental data are presented to validate both the simulator output and the use of mixtures of models for object segmentation or classification. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: [10.1117/1.OE.61.6.063105](https://doi.org/10.1117/1.OE.61.6.063105)]

**Keywords:** lidar; radar; statistics; nuisance alerts; simulator; segmentation.

Paper 20220307G received Mar. 25, 2022; accepted for publication Jun. 6, 2022; published online Jun. 21, 2022.

## 1 Introduction and Related Work

False alarm rates (FAR), constant FAR (CFAR), and receiver operating characteristic (ROC) curves are words more familiar to the radar technology than to the lidar's. For radar, statistical models are commonly used for assessing the validity or degree of confidence of an alert to be issued to a pilot or to a commander. The detection theory, akin to the decision-making theory, has been under development in the radar community since World War II. Reducing the level of nuisance alerts is of utmost importance in the time-pressured aircraft cockpit where decisions must be made rapidly and constantly for the safety of crew or passengers. A document dated 2020 for the radio technical communication for aeronautics reported that there have been several controlled flight into terrain accidents where the terrain awareness and warning systems had been manually inhibited due to the frequent occurrence of nuisance alerts during routine operations.<sup>1,2</sup> In a mission, the level of engagement must rely on information whose degree of confidence must have been assessed somehow. The same can actually be said of the decisions made based on Traffic Collision Avoidance Systems in the air or by air traffic controllers on the ground. The detection theory relies on statistical models of both the radar pulse and the physical environment that it interrogates. Those models in turn assist the engineer in the design of new radar equipment or in the preparation for a mission.<sup>3,4</sup>

Now that it can be said that lidar technology is coming of age and that it is being integrated in advanced driver assistance systems (ADAS) for the automotive industry, in unmanned vehicles of all sorts, and even claims to assist helicopter pilots in the risky task of landing,<sup>5</sup> the same level of rigor is expected from it. The same use of statistical models for FAR, CFAR, and ROC curves has already started its migration to the lidar world over the last 20 years or so, though over a small scale. We found seven papers in our research that explicitly referenced statistical models. In five of them, explicit development of probability distribution functions (PDF) and ROC curves

---

\*Address all correspondence to Robert Bernier, [robert@lidarinnovation.com](mailto:robert@lidarinnovation.com)

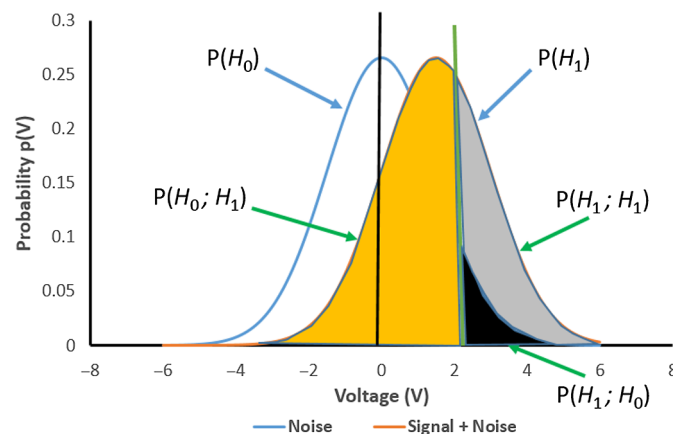
were made in view of predicting instruments' performances, which will be our main goal in this paper. It has been used for the planning and hardware implementation of the data processing to be made on-board the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) mission,<sup>6</sup> where it guided the selection of range dependent thresholds (CFAR) for selecting between the signals from molecules only or from molecules plus particulate contribution. Statistical models have been developed for predicting the effect of rain on the lidar signal in ADAS in the automotive industry,<sup>7</sup> or for the effect of fog on the quality of lidar 3D imagery.<sup>8</sup> A fourth paper used the full statistical approach to predict the performance of a fluorescence lidar in the detection of biowarfare aerosol<sup>9</sup> and a fifth one to evaluate different strategies to mitigate the effect of cross-talk between different signals coming from different vehicles.<sup>10</sup> Two other research papers in the field of lidar only used the concept of ROC curves as a quality metrics for the results obtained from experiments.<sup>11,12</sup> Our team has recently issued a paper in which a statistical model is developed for predicting the effect of snow on the lidar signal, model further used to develop a filtering algorithm to remove the clutter due to snow from the 3D image.<sup>13</sup> This filtering algorithm has been experimentally tested since and we will discuss in this paper the fact that it can be seen as an application of the CFAR technique. Radar and lidar technologies will become the rule for comprehensive data fusion from numerous interconnected vehicles in global intelligent transportation systems<sup>14</sup> and the detection theory must be rigorously used in view of not overloading the data processing systems with nuisance alerts.<sup>15</sup>

In Sec. 2, we will present a brief review of the detection theory essentials. In Sec. 3, we will present our application of this theory to three cases of simulations, one for reporting detections of men overboard in a search and rescue helicopter mission and two for reporting detections of obstacles to a helicopter pilot in the process of landing. In Sec. 4, we will discuss CFAR and we will introduce and demonstrate the possibility of developing mixtures of models, even during a mission, to be used for object segmentation or classification. As far as we know, this will be a novel use of these statistical models. We will then show how accounting for the laser divergence in the simulation of a lidar mission points at 3D scanning lidar statistical models being of a Swerling IV case (a Chi-2 PDF), which again will illustrate how close the lidar and radar technologies are to each other.

## 2 Review of the Detection Theory

The detection theory that will be used is the same as has long been applied in the field of radar, more recently in the field of lidar<sup>16</sup> and also of data fusion.<sup>14</sup> We will briefly present the basic ideas of that theory. We can refer to Appendix A as well as to two important books by Kay<sup>17,18</sup> for the specific mathematical details.

The most important terms are: probability of detection (Pd) and probability of false alarms (Pfa). The relation between Pd and Pfa is shown in Fig. 1, inspired from Kay.<sup>18</sup> We can see two Gaussian PDFs that could represent the voltage of an oscillating signal: the left one represents the



**Fig. 1** Two Gaussian PDFs and two types of errors.

PDF of a white Gaussian noise with zero mean and the right one represents the PDF of the expected “signal plus noise.” In the design of those PDFs, the designer will include all that is known about the physics of the sensor and of the environment it is meant to interrogate. Or, at mission time, the histogram of the already acquired signals could be used to build those PDFs as will be discussed in Sec. 4.

In Fig. 1, the noise only hypothesis (left Gaussian PDF) is called the  $H_0$  hypothesis or null hypothesis. The signal plus noise hypothesis (right Gaussian PDF) is the  $H_1$  hypothesis or alternative hypothesis. The green line represents the threshold placed to decide between the two hypotheses. For simplicity, the signal is represented also by a Gaussian but other PDFs can be used as best fits.

The grayed area under the  $H_1$  hypothesis represents Pd and the black area under the  $H_0$  hypothesis represents Pfa. In most cases of interest, the two hypotheses will be partly overlapping, as shown in Fig. 1, so that increasing Pd by lowering the threshold will increase Pfa. In Fig. 1, the yellow zone at the left of the threshold represents missed detections (false negatives).

As the figure illustrates, Pd and Pfa are the right-tail probabilities (of  $H_1$  or  $H_0$ ): probability of exceeding a certain value [ $Q(x)$  of Appendix A]. Appendix A gives the main equations for the specific case of Gaussian PDFs.

The type I error is to have decided that the hypothesis  $H_1$  was right while the truth was  $H_0$  or  $P(H_1; H_0)$ : a false alarm. The type II error is to have decided that the hypothesis  $H_0$  was right while the truth was  $H_1$  or  $P(H_0; H_1)$ : a missed detection.

Moving the threshold to the right would decrease Pfa but would also increase the probability of missing a detection or, said otherwise, it would decrease Pd. Once a tolerable value has been decided for Pfa, equations are given to determine the value for the threshold. The equations for Gaussian PDFs will be found in Appendix A and Kay.<sup>18</sup>

The important point here is that, if we have *a priori* knowledge about the statistical models of both the  $H_0$  and  $H_1$  hypotheses, we have an objective means for fixing the threshold. The threshold is determined by the rate of false alarms, which is deemed tolerable in view of decreasing the rate of missed detections. The  $H_0$  and  $H_1$  hypotheses may be about any quantifiable feature of the system output. In the simulations of Sec. 3, we will address two distinct features: one is received optical power and the other is height above the ground.

It must be stated here that other PDFs could have been used instead of the Gaussian, and they all lead to associated versions of the  $Q(x)$  functions [Appendix A and Kay<sup>18</sup>]. It is the case of the Swerling IV PDF, which actually is a non-central Chi-2 PDF. It must be noted though that only the Gaussian PDF also leads to its inverse  $Q^{-1}(x)$ . For other PDFs, the designer will have to go the other way around: first, to decide for a certain value of threshold and, second, to see the resulting value for Pfa.

The Neyman–Pearson theorem yields a more objective way to fix the threshold using the likelihood ratio (LR) (Appendix A). The LR will allow to attribute a quantitative level of confidence to each and any of the individual detections. The higher a particular LR is above threshold, the more reliable is the decision that has been made to declare a detection. It could be used either to remove nuisance alerts from a report or to convey the meta-information about the certainty or uncertainty of the report using icons or color charts. This may be a help in the process of decision making.

The relation between Pfa and FAR in a real system under operation would have to account for the pulse repetition frequency (PRF) in the following way:

$$\text{FAR} = \text{Pfa} * \text{PRF}. \quad (1)$$

In a plan position indicator containing  $2000 \times 2000$  points, a Pfa in the order of  $10^{-8}$  would mean that 0.04 false alarms could occur. For a Pfa of  $10^{-4}$ , that number could climb to 400 in the scan.

Finally, the ROC curve is simply a graph of Pd versus the choice of Pfa. The best detector is the one whose Pd climbs the most rapidly with respect to Pfa: the highest probability of detecting targets of interest at the lowest cost in false alarms (see Figs. 4 and 10 of Sec. 3). The best detector is the one that maximizes the area under the ROC curve.

### 3 Modeling Three Lidar Missions

Three simulated lidar missions will be discussed. The simulator has been developed over the last six years for investigating the capabilities of a 3D lidar scanner based on the technology of a double pair of Risley prisms.<sup>19</sup> It can emulate the operation of an airborne 3D lidar scanner over flat ground or over digital elevation models of terrains. Those terrains can be populated with rectangular boxes or cylindrical objects with their associated reflectivities. The user enters a choice of physical parameters for the lidar, including the parameters for the Risley prisms scanner, as well as information about the optical properties (e.g., reflectivity) of the objects in the scene. A definition of the altitude and flying path, including the possibility of it being defined by sequences of latitude, longitude, altitude coordinates, must also be entered by the user. Then, while the vehicle is moving, and upon each laser firing, ray tracing is applied to the laser pulse propagation and interaction with the scene. It is to be noted that the simulator accounts for the possibility of multiple returns, for the interaction with either snow, dust, rain or fog, as well as for the divergence of the laser source. It also includes an algorithm for a first stage of segmentation in the points cloud. The results of the calculations are graphically overlaid on the scene.

#### 3.1 Simulating a Search and Rescue at Sea Mission

This simulation has led to a typical application of the statistical model: to determine the appropriate threshold value for the detector output. Figure 2 shows a scenario for a helicopter flying at an altitude of 500 m. The red dots are for models of retro-reflectors. Their size is made equal to the diameter of the laser beam at their respective range from the sensor when the sensor passes above the retro-reflector at angle 0 deg. The green dots are for targets of reflectivity 1. The blue dots for targets with reflectivity 0.3. Both green and blue targets have size 1 m × 1 m. All targets extend up to 20 cm above the sea surface. The sea is represented as a flat surface (no waves) with reflectivity 0.02 because a low water reflectivity is characteristic of a lidar operating at the eye-safe 1.56- $\mu\text{m}$  wavelength. The helicopter is flying at the constant speed of 100 knots. The simulated lidar is based on one single pair of Risley prisms with a 45-deg field-of-view (FOV) and is kept pointing at nadir.

The theoretical values of Pfa and Pd have been calculated with Eqs. (5) and (6) of Appendix A at the different values of threshold to be discussed. The calculation of the theoretical Pfa is based on a Gaussian model for the return from water ( $H_0$  hypothesis) with mean power 6 nW and standard deviation 15 nW due to the addition of a zero mean Gaussian noise. The theoretical values of Pd are based on Gaussian models for the return from targets ( $H_1$  hypotheses) with mean powers from 10 to 200 nW and standard deviation 15 nW due to the same noise being added. Implicitly, no dispersion of the optical powers was designed to be due to the targets

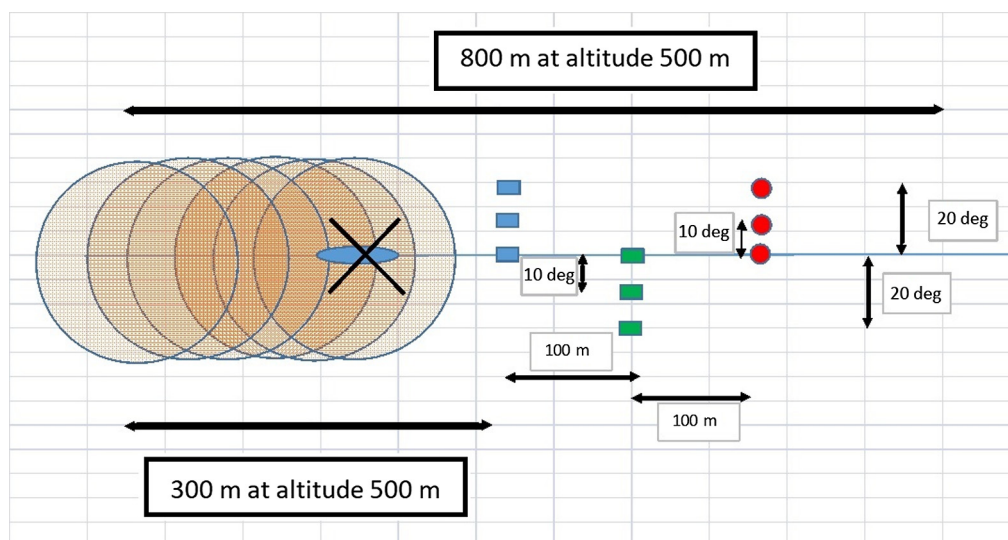
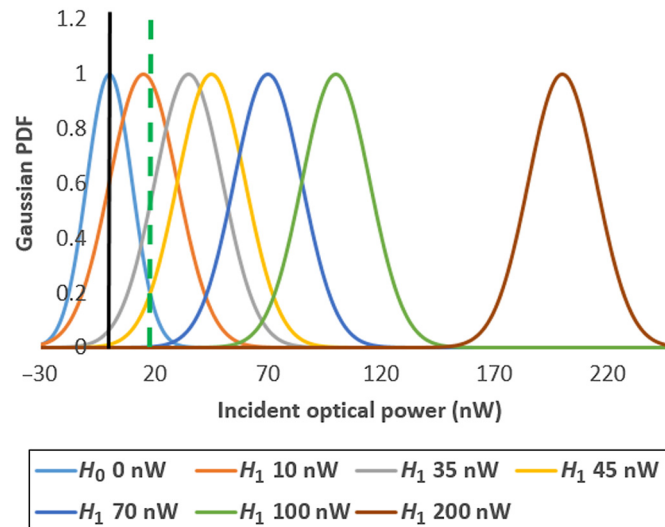


Fig. 2 Illustration of the parameters for a search and rescue at sea mission.



**Fig. 3** PDFs for the  $H_0$  and various  $H_1$  hypotheses for the search and rescue at sea mission.

themselves, only to the noise. We will discuss later in Sec. 4 that this is not accurate in regards to the real response of the targets. The Gaussian PDFs for the  $H_0$  and the multiple  $H_1$  hypotheses are shown in Fig. 3. It is worth noting that out of the actual lidar electronic detection chain, no negative signal should occur.

In the figure, we have arbitrarily set a threshold (green dashed vertical line). In Fig. 3, we can see that the 10-, 35-, and 45-nW  $H_1$  hypotheses for the targets have some overlapping with the  $H_0$  noise hypothesis centered at 0 nW. For those, there will be a non-null FAR. But for the other  $H_1$  hypotheses, the FAR should be very low, even null, being given the location of the threshold.

In Table 1, we show the theoretical values of Pd for each  $H_1$  hypothesis being given six different values of Pfa corresponding to six pre-selected values of threshold. These were calculated using Eqs. (5) and (6) of Appendix A.

In Table 1, we see that, for each value of Pfa, there corresponds only one value of threshold. However, for those values of Pfa and threshold set in accordance with the model for the noise, we see that there correspond different values of Pd, one for each  $H_1$  hypothesis since their respective distance away from the threshold varies as illustrated in Fig. 3. The numbers in the column “simulation results” are actual values of Pd and will be explained later, along with the numbers in Table 2.

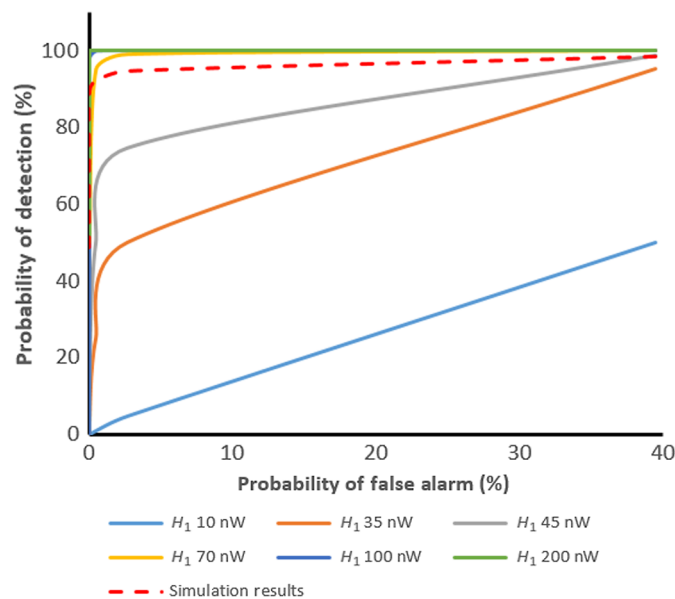
The calculated ROC curves, one for each  $H_1$  hypothesis, are shown in Fig. 4. In it, we see that the ROC curve always remains very low for the  $H_1$  10-nW hypothesis. This happens because the lowest value of threshold (10 nW) for the highest value of Pfa (39.49%) already leaves behind, as missed detections, half of its Gaussian PDF centered at 10 nW. These ROC curves are very representative of what the mission designer may expect as system performance but only if all

**Table 1** Values of Pd (%) for all  $H_1$  models for six different values of threshold and Pfa.

Pfa (%)	Threshold (nW)	$H_1 - H_{10}$	$H_1 - H_{35}$	$H_1 - H_{45}$	$H_1 - H_{70}$	$H_1 - H_{100}$	$H_1 - H_{200}$	Simulation results
39.49	10	50	95.2	99	100	100	100	98.7
2.66	35	4.8	50	74.8	99	100	100	94.9
0.47	45	1	25.2	50	95.2	99.9	100	92.3
9.9E-4	70	3.2E-3	1	4.8	50	97.7	100	88.5
1.8E-8	100	9.9E-8	7.3E-4	0.01	2.3	50	100	79.5
0	200	0	0	0	0	1.3E-9	50	48.7

**Table 2** Simulation results at altitude 500 m.

Pfa (%)	Threshold (nW)	Points from Water	Points from Targets	FAR (%)	Det. rate (%)	Mis. rate (%)
N/A	0.01	1.6E+6	78	N/A	N/A	N/A
39.49	10	6.5E+5	77	40.3	98.7	1.3
2.66	35	4.8E+4	74	2.95	94.9	5.1
0.47	45	8.7E+3	72	0.54	92.3	7.7
9.9E-4	70	30	69	1.9E-3	88.5	11.5
1.8E-8	100	3	62	1.9E-4	79.5	20.5
0	200	2	38	1.2E-4	48.7	51.3

**Fig. 4** ROC curves for all  $H_1$  hypotheses for the search and rescue mission at 500 m.

targets correspond to the  $H_1$  10 nW or to the  $H_1$  35 nW or to only one of the other various  $H_1$  hypotheses. In the figure, a curve has been added: it displays the actual results of the simulations. It corresponds to none of the theoretical predictions, though closely to the  $H_1$  70 nW, and we will proceed to explain how it has been obtained. It corresponds to the numbers in the last column of Table 1.

If no or a very low threshold is set (0.01 nW in Table 2), during the full flight, the simulation reports  $1.63 \times 10^6$  points from the sea surface and 78 points from targets: all possible returns are recorded as raw data before the detection process. To obtain the numbers shown under the title simulation results in Table 1, we counted how many of the 78 target points were recorded for each level of threshold being set in the raw data. Table 2 gives the results of the counting procedure.

All calculation examples to follow refer to the 10-nW threshold line in Table 2. The FAR has been counted as the number of alarms due to the sea surface (653,277 returns/1,623,183 possible returns from sea = 40.25%). The detection rate has been counted as true positives from targets (77 returns/78 possible returns = 98.7%). The missing rate has been counted as missing target points (1 miss/78 possible returns = 1.28%). FAR is the level of false alarms (nuisance) when water was reported above threshold as if a target. We see from Table 2 that, if we let Pfa (FAR) go up, so also does Pd (Det. Rate), whereas missing rate goes down: this corresponds to pulling the

threshold to the left in Fig. 3. We can also see that the FAR (%) values out of the simulator are very close to the theoretically predicted ones in the first column. Finally, we note that by increasing the threshold level above 45 nW, we remove almost all nuisance alerts without losing too many points on the targets. The same target can be hit multiple times, from different ranges and angles, while the helicopter is moving.

In Fig. 4, the fact that the experimental ROC curve shows higher detection rates than most of the theoretical ROC curves will be explained in Sec. 4 by the fact that most targets actually had a non Gaussian PDF with large dispersion in the levels of power returned, the highest returns thus led to rates of detection higher than predicted. In Sec. 4.2, a histogram of experimental data will show that such behavior of the targets is not an artifact due to our simulator but can be obtained in real situations.

Hence, a threshold set at 45 nW would be a good choice while preparing the mission if we have documented confidence in the statistical models we have at hand both for the sea reflectivity (noise) and for the targets of interest. We will discuss and demonstrate in Sec. 4 how processing of raw data histograms could have led to setting the threshold properly at mission time when we have no such *a priori* information. The new generation of lidars with full waveform recording and on-board FPGA data processing allows that to be done.

### 3.2 Simulating a Helicopter Landing

The aim was to provide information to a helicopter pilot concerning the safety of a landing zone (LZ). The LZ was to be 20 m  $\times$  20 m large and, in it, six posts of dimensions 5 cm  $\times$  5 cm  $\times$  1 m and reflectivity 0.3 were randomly distributed. The roughness of the terrain was simulated by adding on the ground a randomly generated “height” noise. Two cases of roughness were tested: 0 to 5 cm and 0 to 10 cm. The posts would emerge out of the roughness. The reflectivity of the terrain was also set to 0.3. Figure 5 shows the scene. The helicopter motion is modeled as coming along the Y axis direction, from a distance of 1200 m, a starting height of 500 m and a descending rate of 152 m/min (500 ft/min). The starting speed is 100 knots at distance 1200 m from the LZ and is uniformly decreasing.

Two scenarios were tested: one with a generic (GENR in Fig. 6) 3D lidar scanner based on a single pair of Risley prisms, the other with a double pair of Risley prisms (DRP in Fig. 7). As described in our 2015 paper,<sup>19</sup> in the DRP, the inner pair rotates and samples a high resolution FOV of 30 deg. The outer pair rotation (90-deg FOV) is constantly adjusted in order to keep the inner pair FOV pointing at the LZ while the helicopter approaches. For the case of GENR, the LZ

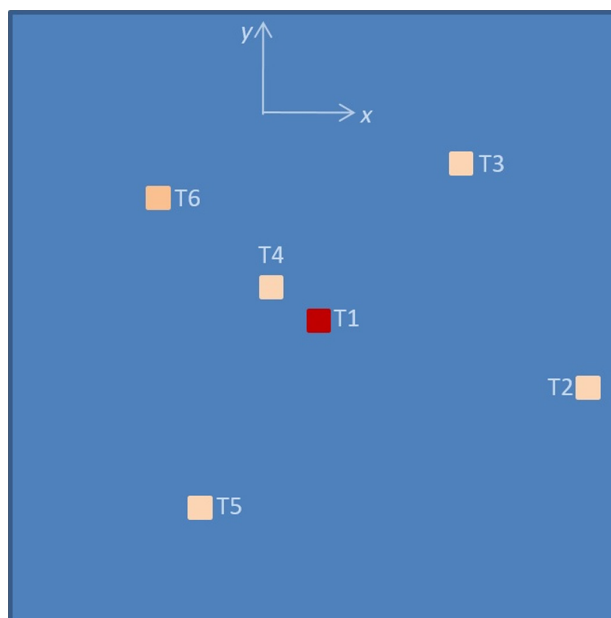
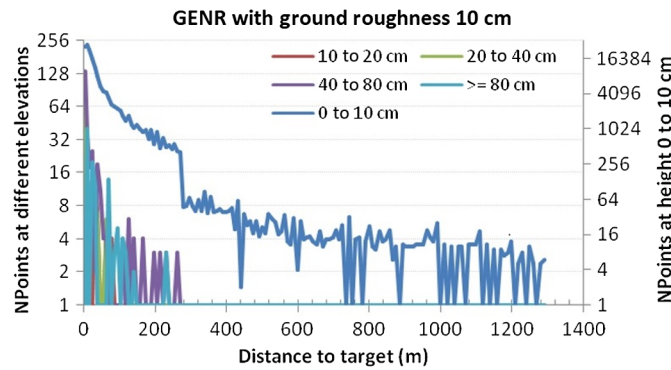
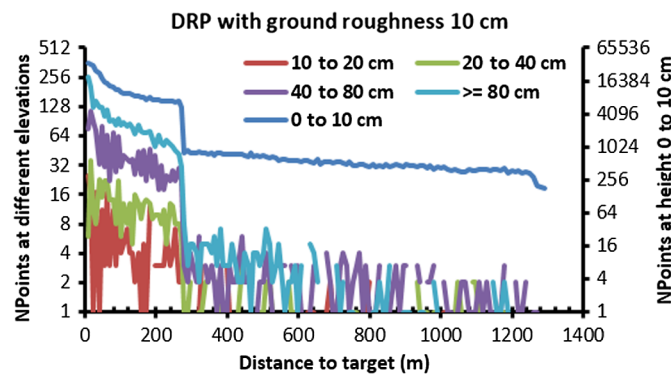


Fig. 5 Schematics of the LZ with obstacles.



**Fig. 6** Performance of a GENR lidar scanner for sampling the LZ as a function of distance. Numbers of detections at various heights on the posts from all ranges.



**Fig. 7** Performance of a DRP lidar scanner for sampling the LZ as a function of distance. Numbers of detections at various heights on the posts from all ranges.

is within its FOV only under some circumstances during the mission while, for the DRP, the LZ is constantly being sampled. Figure 6 shows the number of points detected in the LZ with the GENR as a function of distance for a ground roughness of 10 cm. Figure 7 does the same for the DRP. The sharp drop around 300 m is due to the laser PRF passing from 200 kHz down to 25 kHz, with higher energy, at larger distances from the LZ.

In Figs. 6 and 7, the blue line with legend 0 to 10 cm gives the number of points detected on the ground within the LZ (right vertical scale in the figures); 10 to 20 cm, the points detected on posts between 10 and 20 cm and so on for the other lines. The DRP scheme of scanning promises an improved job at starting sampling the LZ and mainly the obstacles in the LZ from much further away, thus giving the pilot more time for decision making.

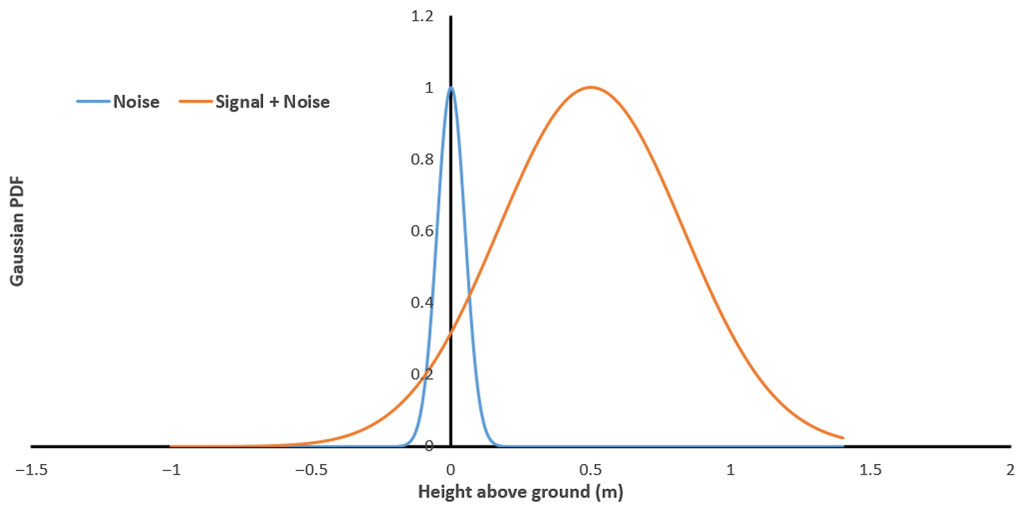
The threshold to be fixed in this application was a height threshold in a point cloud in which local models of the ground were first developed, following a random sample consensus approach (RANSAC), by fitting planes of different heights and angles, which would best match the roughness locally. The aim is now to decide on the presence of obstacles above the ground rather than on optical power thresholds.

The model for the  $H_0$  hypothesis was a Gaussian with mean 0 cm and standard deviations 2.5 and 5 cm for the 5 and 10 cm roughness cases, respectively. The model for the  $H_1$  hypothesis was a Gaussian with mean 50 cm and standard deviation 33 cm. This is shown in Fig. 8 for a terrain roughness of 10 cm.

Table 3 gives the selection of values for Pfa and the thresholds that were calculated from them with Eqs. (5) and (6) of Appendix A.

Figure 9 shows the false alarms that were generated (vertical axis) when compared to those predicted (horizontal axis).

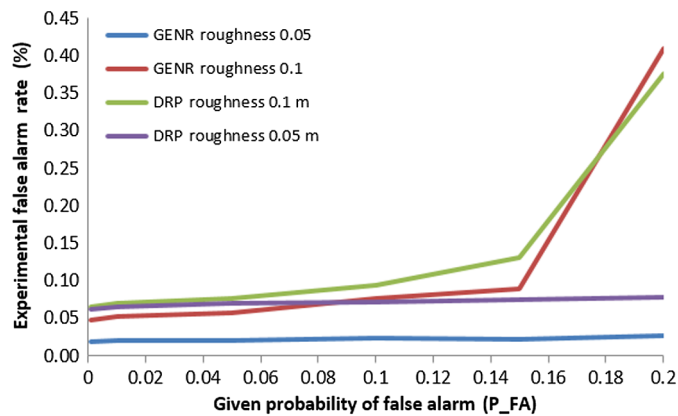




**Fig. 8** Illustration of the Gaussian PDFs for the heights of the terrain and the obstacles in the LZ.

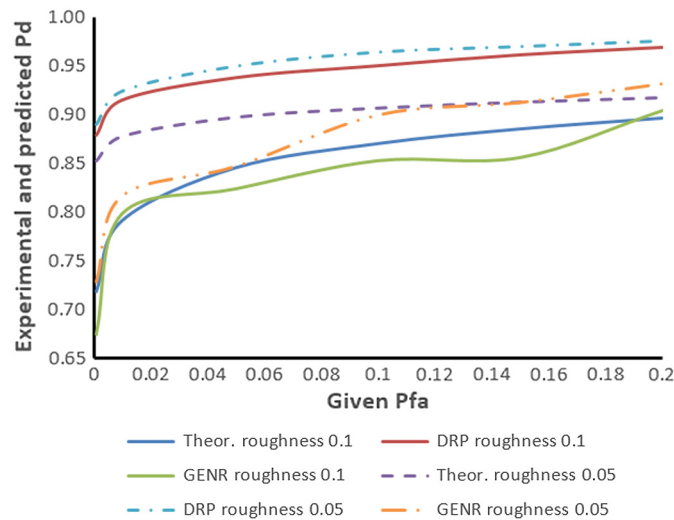
**Table 3** Selected values of Pfa and the thresholds (m) for the landing scenario.

Pfa	5-cm roughness (m)	10-cm roughness (m)
0.001	0.1545	0.309
0.01	0.1163	0.2326
0.05	0.0822	0.1645
0.1	0.0641	0.1282
0.15	0.0518	0.1036
0.2	0.0421	0.0842



**Fig. 9** False alarms reported versus false alarms predicted during helicopter landing simulation.

If the model had predicted correctly, the experimental false alarms should have made straight lines at 45 deg with respect to the horizontal axis. Instead, we see that the FARs for both the GENR and the DRP at the 5-cm roughness remain almost constant and low. To explain that, we need to return to the definition of Pfa as given in Eq. (5) in Appendix A: it is the right-tail probability for the hypothesis  $H_0$ . It does give a maximal value, which will be realized the more



**Fig. 10** Experimental and theoretical ROC curves for the simulation of a landing helicopter.

the two hypotheses are overlapping. With the 5-cm roughness terrain, this is less the case as both hypotheses are more separated. In the case of the 10-cm roughness, the very lowest parts of the  $H_1$  hypothesis come in more overlap with the highest thresholds of the hypothesis  $H_0$ .

Last, it is important to note that, in the analysis of data, we found effective values of LR anywhere between 2 and 1000, thus much above the threshold: in a real mission, this LR value can be used either to remove nuisance alerts from a report or to convey the meta-information about the certainty or uncertainty of the alerts.

Figure 10 shows the experimental and theoretical ROC curves for the scenarios of helicopter landing.

In Fig. 10, we observe that the predicted ROC curve is always closer to the experimental one for the GENS case than for the DRP case. When looking at the details of the detections, it was observed that the mean of the heights of the points detected on the posts is 0.497 m for the case of GENS while it is 0.741 m for the case of the DRP. In the case of the DRP, the detection is dominated by the two targets at the very center of the LZ (T1 and T4 in Fig. 5), where the scanner is tracking, and those two targets yield heights close to or above 0.9 m toward the end of the scan, when the helicopter is almost hovering above them. Such highly elevated points are further away from the model of the ground, thus yielding an enhanced detection rate. The use of simulations along with the statistical analysis of results allow a better mission preparation.

ROC curves can be a powerful tool both for designing a lidar instrument and for preparing a mission. By inserting in the statistical models all knowledge about the environment and objects to be interrogated as well as all the design parameters to be tested, the engineer can get a quick grasp at what best choice to make. The best choice is the one that maximizes the area under the ROC curve. This is what was done in some of the Refs. 6–12. For the cases presented here, in Figs. 4 and 10, we have been able to explain the differences between the theoretical predictions and the results of simulation via imperfections in the models for the targets. We will discuss this topic in Sec. 4 and a possible remedy to that in our Conclusion.

## 4 Discussion

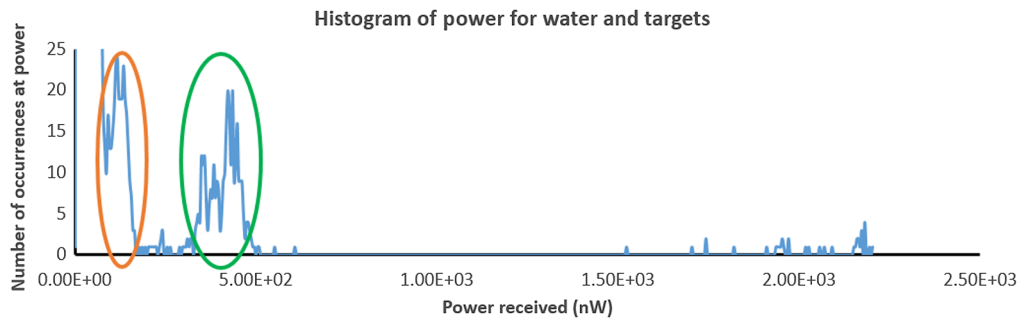
The theory and results presented above let us foresee at least two problems: (1) what can we do if we have no *a priori* knowledge of the scene and (2) what can we do if the threshold to be set varies in the scene? The discussion of the first problem will have us introduce the idea of mixtures of models while, along the way, discovering that Gaussian PDFs may not always be the best statistical models for a 3D scanning lidar. Discussion of the second problem will allow us to introduce briefly the concept of CFAR.

#### 4.1 Why and How a Mixture of Models?

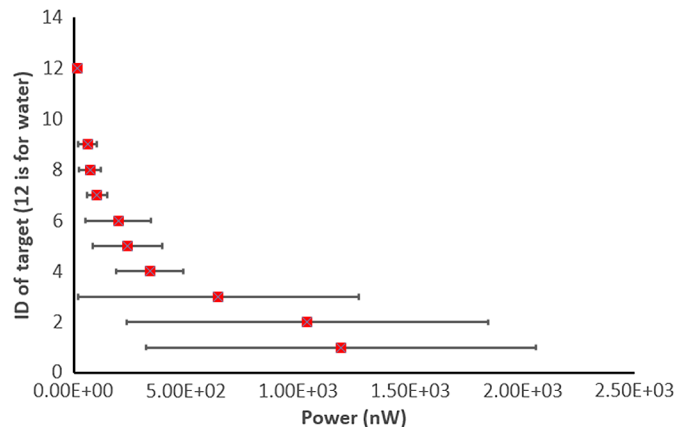
The novel idea to be put forward and developed in this section is an adaptation of the Gaussian decomposition procedure used by Wagner et al<sup>20</sup> or Li<sup>21</sup> for retrieving individual lidar returns under an enlarged envelope signal. This happens when, for instance, a lidar laser pulse encounters successive closely packed reflections in a tree foliage. The idea is to find, under the envelope signal, the number, location, and width of Gaussian shaped laser pulses whose additive superposition best reproduces the envelope signal, thus yielding spatial super-resolution in the data.

Making an analogy between a lidar waveform and the histogram of raw data, each cluster of occurrences in the histogram could be seen as the image of one specific statistical model, thus eventually representing one class of objects. Raw data acquired during a mission could be thus analyzed and be used (1) for setting an appropriate threshold or even (2) for starting a process of segmentation and classification in a scene. The abundant literature in segmentation shows that there is no one-fits-all tool. The procedure presented here would just add one more tool. We first use the setup of Sec. 3.1 to illustrate the method with the helicopter speed reduced from 100 to 20 knots.

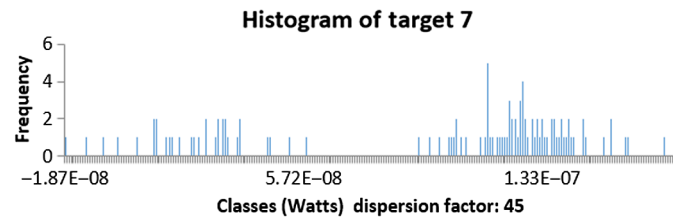
Figure 11 shows the histogram of the raw data of the scenario at altitude 500 m. Figure 12 shows the powers returned by each target or by water, and their dispersion. From Fig. 12, it can be said that the targets 7, 8, and 9 (blue targets with reflectivity 0.3 in Fig. 2) are accounted for by the peak, which has been surrounded in orange in the histogram of Fig. 11; however, some of their returns are under the large peak for water. The targets 4, 5, and 6 (green with reflectivity 1 in Fig. 2) have their values mostly under the peak surrounded in green; however, many of their values also go under the orange peak while low values for targets 1, 2, and 3 (red retro-reflectors in Fig. 2) are also under that same green peak.



**Fig. 11** Histogram of the raw data of the search and rescue at sea mission (altitude 500 m).



**Fig. 12** Illustration of the dispersion in the power received from each target and from water during the full mission at altitude 500 m (nW).



**Fig. 13** Histogram of power returned by target 7 at 0 degree (in Watts).

#### 4.2 Which Statistical Model for a Scanning Lidar?

The concept is thus to try to reproduce the histogram of Fig. 11 with a mixture of statistical models. However, before choosing one type of model, we need to take a closer look at the histograms of the individual targets. Figures 13 and 14 show some of those at 0 deg and at maximal 20 deg: they are representative of all nine targets.

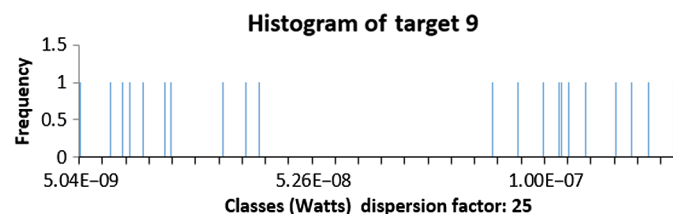
In all the nine histograms, a clear cut is observed between low level returns and higher level returns with dispersion factors going from a minimum at 9.8 to a maximum at 313 (45 and 25 for targets 7 and 9 in Figs. 13 and 14, respectively). During the full mission while the helicopter is first approaching and then receding from a target, looking at some from angles varying between 0 and  $\pm 22.5$  deg, the dispersion of power due to the varying distances and angles could account for a maximum dispersion factor of 3 approximately. Something else must cause such large values of dispersion.

Our simulator does account for the divergence of the laser beam. The beam is constructed as one central ray comprising 50% of the energy surrounded by eight peripheral rays each carrying 6.25% of the energy. Our analysis of this feature has shown that, when the laser beam hits a target, it may be sometimes with only the central beam plus one peripheral beam and sometimes by only one peripheral beam. Another part of the beam may hit the water. This effect can also be seen in real experimental data and may have as a result a distortion of the objects 3D shapes.

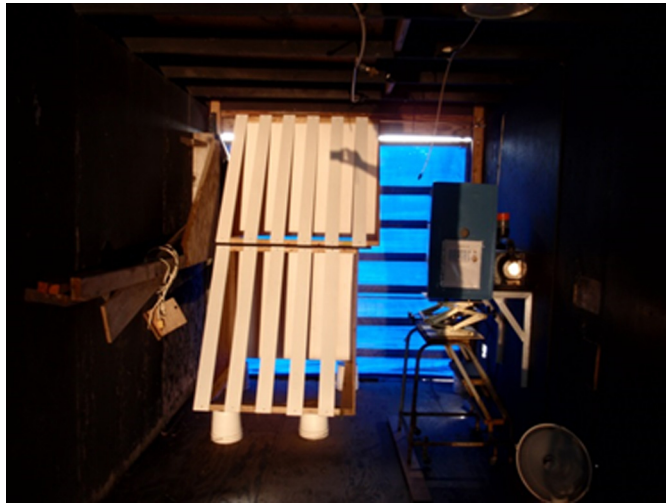
In an experiment set up rightly to evaluate this effect of the laser divergence,<sup>22</sup> a 3D target was installed in an aerosol chamber. As seen in Fig. 15, the target is made out of white painted wooden planks laid at 30 deg in front of a vertical white wooden board. The front planks and the intervals between them are 10-cm wide. The lidar equipment used to scan the target was a Lumibird OPAL3 with a 45-deg full FOV.

Figure 16 is a typical histogram of the range corrected intensities collected along any of the front planks: range correcting the intensities means multiplying them by the square of their range from the sensor. In it, we see the same split of intensity values as reproduced by our simulator, showing that this effect is not an artifact of the simulator. The splitting is caused by some fraction only of at least some laser beams hitting the front plank and the other part hitting the back board. It is important to note that, due to this split in intensity values, the actual working of the electronics or software process in view of finding the range at which the threshold is crossed may produce a jitter in the range bin location of the plank, thus inducing a distortion of the 3D shape.

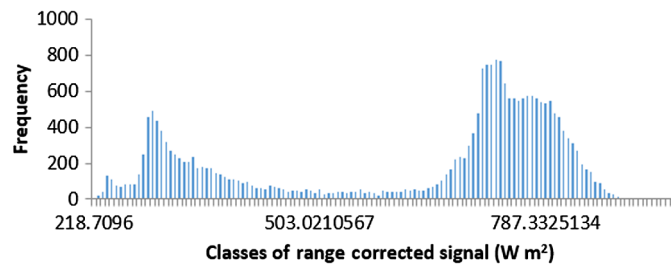
Various statistical models were tested for matching the cumulative histogram of raw data for each target in Fig. 2, including a Gaussian, a Rician, and a Swerling IV. In the case of radars, the Swerling models are used to describe the statistics of fluctuating targets, the fluctuations being caused by the variation of the target cross section caused by the variation of the target aspect.



**Fig. 14** Histogram of power returned by target 9 at 20 deg (in Watts).

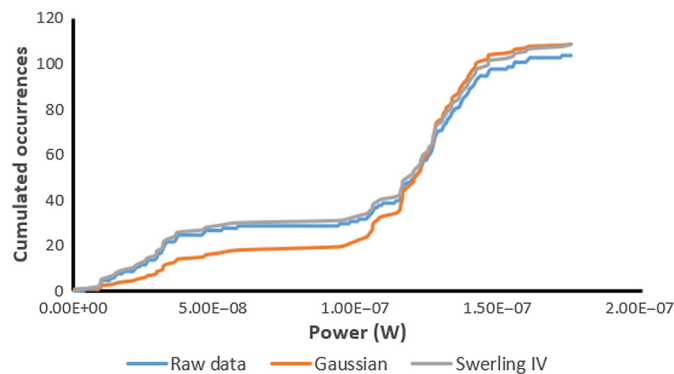


**Fig. 15** White triangular shape 3D target in aerosol chamber.

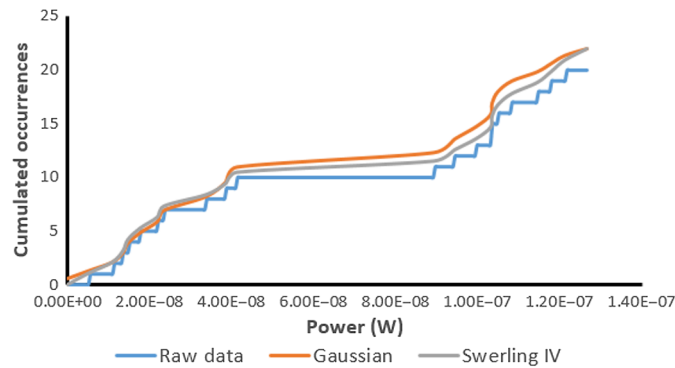


**Fig. 16** Histogram of power returned from the 3D target of Fig. 15.

Swerling models are instances of Chi-2 PDFs. The Swerling IV model is for a target having one main scattering element that predominates together with many smaller independent scattering elements.<sup>4</sup> This model could reproduce the case where either the full ray or only the central ray or only a peripheral ray would impact the target at different incident angles. In Figs. 17 and 18, we show the results of the Gaussian and of the Swerling IV models, which were the only ones close to matching the raw data cumulative histograms and only for the same targets as for Figs. 13 and 14, but the results are representative of all nine targets of the simulation.



**Fig. 17** Cumulative histograms of Gaussian and Swerling IV statistical models with cumulative histogram of raw data for target 7 at 0 deg.



**Fig. 18** Cumulative histograms of Gaussian and Swerling IV statistical models with cumulative histogram of raw data for target 9 at 20 deg.

For the Gaussian model, we used the standard equation [Ref. 18, Eq. (2.1), p. 20]

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right), \quad (2)$$

where  $\mu$  and  $\sigma$  are the mean and standard deviation of the powers returned by the target, respectively. For the Swerling IV model, we used the Chi-2 (non-central) equation [Ref. 18, Eq. (2.15), p. 26]

$$p(x) = \frac{x^{\left(\frac{\nu}{2}-1\right)} \exp\left(-\frac{1}{2}(x + \lambda)\right)}{2^{\frac{\nu}{2}}} \sum_{k=0}^{k=N} \frac{\left(\frac{\lambda x}{4}\right)^k}{k! \Gamma\left(\frac{\nu}{2} + k\right)}, \quad (3)$$

where  $\nu$  is the number of degrees of freedom (here  $\nu = 2$ ),  $\lambda$  is the sum of the square of the mean of the low values and the square of the mean of the high values of power for the target, and  $\Gamma$  is the Gamma function and we limited  $k$  at  $N = 10$ .

In Figs. 17 and 18, we see how well related to a scanning radar a 3D scanning lidar may be: the Swerling IV models always better reproduce the cumulative histograms. This is one more reason for integrating in the lidar technology the same general statistical approaches, which have been customary to the world of radar for decades now.

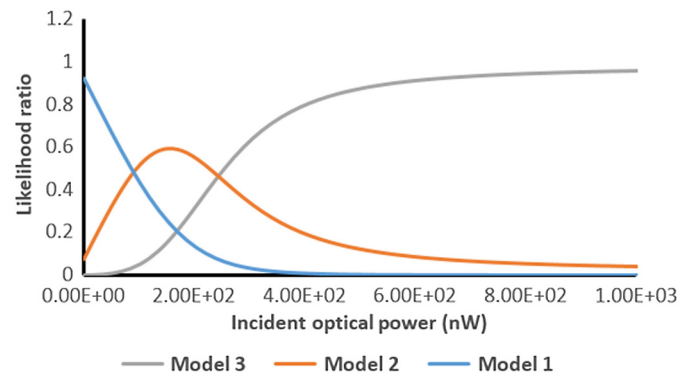
### 4.3 Mixture of Models and Its Use for Segmentation/Classification

In its approach to Gaussian decomposition, Li<sup>21</sup> adapted the algorithms for unsupervised learning developed by Oliver et al.<sup>23</sup> We ourselves adapted those equations to a situation where Gaussian models are replaced by Swerling IV models. The equation we used is as follows:

$$Q_{ij} = \frac{p_j * f_j(x_i)}{\sum_{j=1}^{j=N} p_j * f_j(x_i)}. \quad (4)$$

In Eq. (4),  $p_j$  is the probability associated to model number  $j$ ,  $f_j(x_i)$  is the value of Eq. (3) for the Swerling IV model number  $j$ , with its associated value for  $\lambda$ , at the power  $x_i$  associated to the bin number  $i$  in the histogram.  $N$  is the number of models (number of clusters of occurrences found in the histogram). A peak searching algorithm found the number and location of peaks in the histogram of Fig. 11 and arrived at  $N = 6$  models. The value of  $p_j$  was arbitrarily set at  $1/N$  for each model.

$Q_{ij}$  is the likelihood for the specific power  $x_i$  to belong to the statistical model number  $j$ .<sup>20,21</sup> In Fig. 19, we show the results of applying Eq. (4) to the histogram of all raw data out of the simulator for the search and rescue at sea mission at altitude 500 m. Though six models were found by the peak searching algorithm, only the first three are shown since the last three were more than 10 orders of magnitude lower in their likelihood.

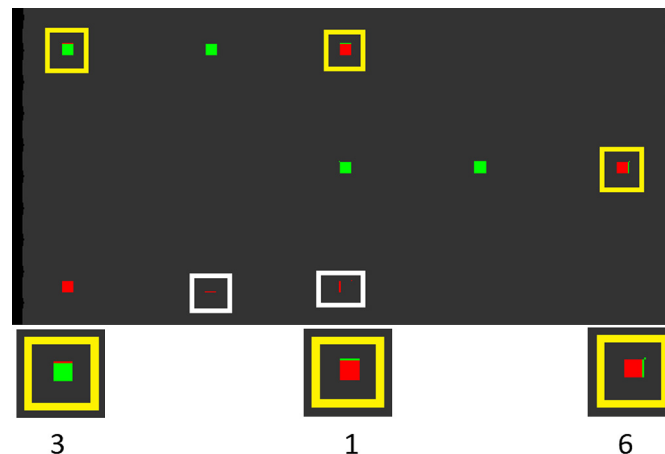


**Fig. 19** Mixture of three Swerling IV models for the histogram of the raw data from the search and rescue at sea mission at altitude 500 m.

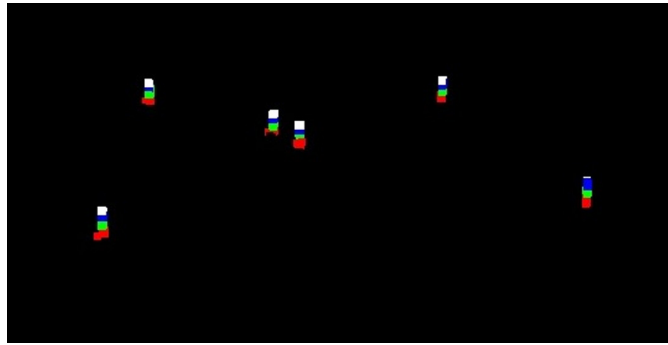
If the meaning of  $Q_{ij}$  is really as written above, then anytime any return power is found in a scan, it should be associated to the model that shows the largest likelihood for that power in our Fig. 19. These three models here could thus be said to define classes of objects interrogated by the lidar. Referring to Fig. 12, we can think model 1 of Fig. 19 to represent returns from the sea surface as well as low levels in Targets 7, 8, and 9. Model 2 would represent the targets 4, 5, and 6 as well as higher returns from targets 7, 8, and 9, and model 3 would represent mainly targets 1, 2, and 3. We have applied this classification scheme in our simulator and the results are shown in Fig. 20. In Fig. 20, model 1 (sea) is painted in dark gray, model 2 in red, and model 3 in green (the detected points have been pictorially enlarged for better visual inspection by the reader). In Fig. 20, all “nuisance alerts” of model 1 are painted in gray but they have rightly been rejected by the classification/segmentation process, which thus acted as a threshold setting. In a real display for pilot, they would have been painted in black for the pilot to concentrate on the targets only. They are shown in gray in Fig. 20 to illustrate that they have been recorded by the sensor but rejected by the method.

In Fig. 20, two white rectangles surround the targets 7 and 8 to show they are fading. Since the dispersion of power was great in all targets (Fig. 12), the targets could appear in one scan and disappear from another (missed detection) or upgrade from model 2 to model 3. Also, three yellow rectangles drive our attention to the fact that parts of targets 1, 3, and 6 have been associated to model 3, whereas other parts of them had been associated to model 2 in the previous scans, or vice versa going from 2 to 3. Enlarged views of these detections are shown at the bottom of Fig. 20, again for ease of visual inspection.

In Sec. 3.1, we described the simulation for the preparation of a search and rescue at sea mission. With good *a priori* models of what there will be to look for, the appropriate setting



**Fig. 20** Classification/segmentation of the raw data from the search and rescue at sea mission at altitude 500 m using three Swerling IV models with model 1 (mainly sea) in dark gray.



**Fig. 21** Classification/segmentation of the raw data from the two scenarios of landing helicopter: scenario DRP.

for the threshold can be fixed beforehand. When no *a priori* knowledge is available, we can think of the process just described as something that could be done during the mission as the new generation of lidars allow. First, one pass over a region of interest could be done to gather raw data for a histogram and produce the models mixture. It would be followed by a second pass with the appropriate thresholding or classification, the classification deduced from the statistical models built out of the raw data from the first pass.

The method was also applied to the results of the landing helicopter of Sec. 3.2. Gaussian PDFs (not Swerling IV) were used to represent the distributions of heights of hits on posts. For the case of the DRP, where Fig. 7 showed that many more points were detected in the LZ than for the case of the GENR of Fig. 6, the method found 6 different models for heights, each represented by different colors in Fig. 21. For the case of GENR, with less points in the LZ, the method found only two models: one for the ground (black in Fig. 22) and the other for the targets (white in Fig. 22). The method however conducted to properly segmenting all posts in the LZ in both scenarios.

Many segmentation methods have been developed over the years. Our knowledge and hands-on experience of many of them is that most require some degree of human supervision as well as much computing power. The results we have displayed in the Sec. 4.3 have been obtained with basic mathematical analysis and computing methods, which can be achieved quickly at mission time in an unsupervised manner, which may prove a great advantage. We show at the end of Sec. 4.4 its application to real data in a complex scene degraded by snow.

#### 4.4 Case of CFAR

We now address our second problem: what can we do if the threshold to be set should vary in the scene? We discuss the topic of CFAR, which is simply the result of letting the threshold vary in the scene in order to keep constant the value of FAR everywhere in it. This method is widely used in the field of radar often to compensate for large signals at close ranges due to antennae side-lobes. The technique has also been used for the CALIPSO lidar mission.<sup>6</sup>



**Fig. 22** Classification/segmentation of the raw data from the two scenarios of landing helicopter: scenario GENR.



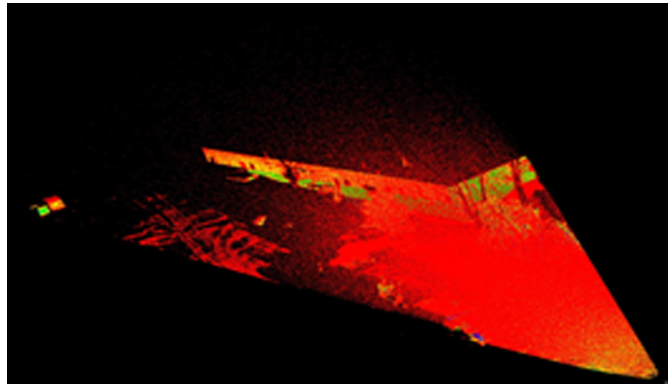
In our work on the physical model of snow,<sup>13</sup> we have developed a method of filtering the signal from snow. What we are doing is to reject any return with an intensity lower than would have been the case for a 0.02 reflectivity target at the same range. This is not different from raising the detection threshold so that these returns will now be rejected, whatever their range from the sensor.

The filter method has since been tested in a real experimental setup. Targets were set up at distances about 56 and 58 m from the sensor. Three targets with known reflectivities were used as reference as shown in Fig. 23. The tests were performed in the parking lot of Lumibird Canada, formerly Neptec Technology Corporation in Kanata Ontario. The Lumibird lidar equipment OPAL3 with a 45-deg full FOV was used. A calibration of the values of intensities of the instrument in terms of incident optical power was used, in the standard lidar equation, to compare all returns from all ranges to the return from a hypothetical target of reflectivity 0.02 at the same range: all intensities below reflectivity 0.02 were filtered out.

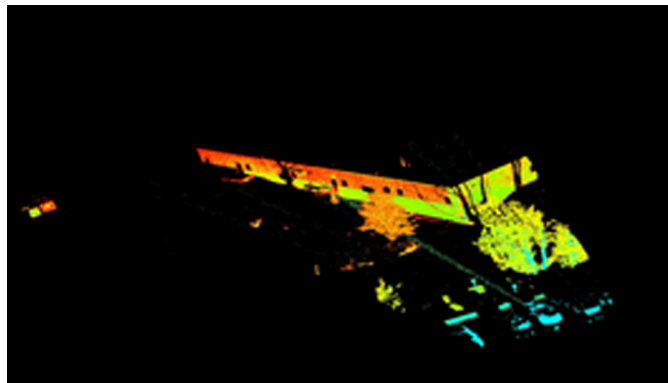
Figures 24–26 show the results: in Fig. 24, the unfiltered image with a red cone due to snow around the sensor; in Fig. 25, the filtered image; in Fig. 26, a rendering of all the points having been correctly or incorrectly removed by the filter. Figure 25 shows a marked improvement due to filtering. The application of the filter results in cars and trees now appearing in Fig. 25 (bottom right), which were buried under the snow returns in Fig. 24. Figure 26 however shows that information has been lost on the ground at the left of the building and on the building itself. This would have been a nice case of applying a CFAR approach. Our physical model of snow shows that the effect of snow decreases rapidly with range. As a consequence, instead of raising the threshold for all ranges, we should have raised it only for a few tens of meters in front of the sensor or, better still, making it high close to the sensor and decreasing it as a function of range in accordance with our physical model. Lidars providing multiple returns information, as is the case of the OPAL3 for instance, could approximate such a filter by simply removing the nearest returns or processing them differently.



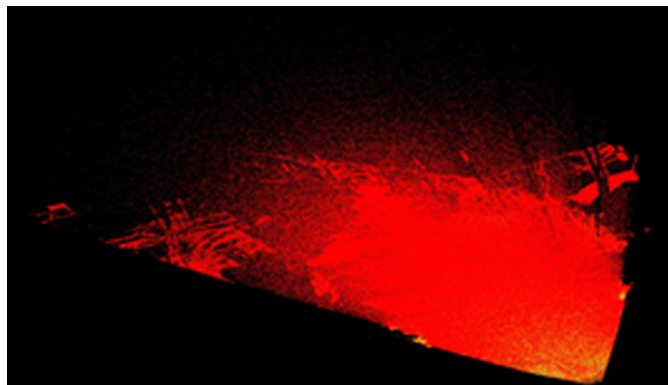
**Fig. 23** Experimental setup for testing a filter against snow.



**Fig. 24** Snow filtering model experimental verification: unfiltered image with red cone due to snow.



**Fig. 25** Snow filtering model experimental verification: filtered image.



**Fig. 26** Snow filtering model experimental verification: showing the points removed by the filter including points on solid targets.

The segmentation/classification method of Sec. 4.3 has been applied there on simulation data only. We will use the points cloud of Fig. 24 to show that the same method can be usefully applied to real experimental data. The intensities in the points cloud of Fig. 24 have first been range corrected.

The histogram of these range corrected intensities was produced and the six Gaussian models calculated from it with Eq. (4) are shown in Fig. 27.

If images of only one class at a time were produced, the user could scroll down through at least six different images and look for objects of interest. For instance, in Fig. 28, we show the points cloud of the scene when the points associated to class 1 (mainly snow) are removed from

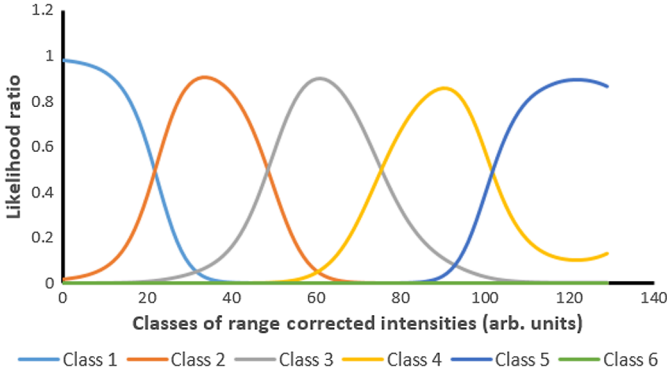


Fig. 27 Mixture of six Gaussian models for the histogram of the range corrected version of Fig. 24.

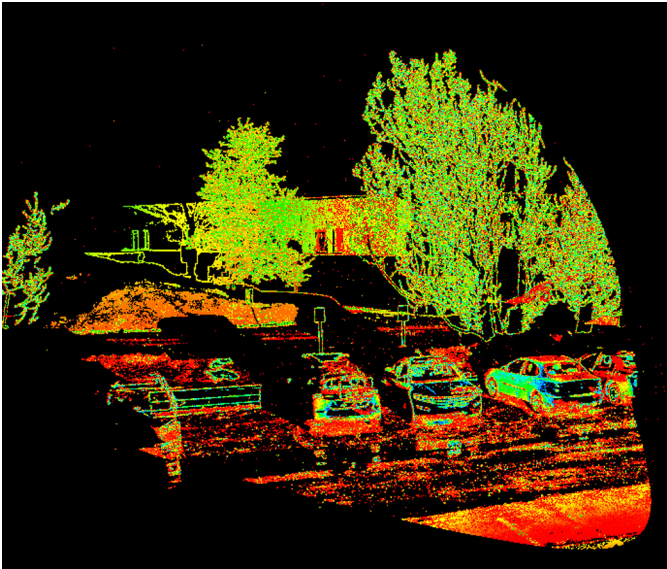


Fig. 28 Scene with points of class 1 (mainly snow) being removed.

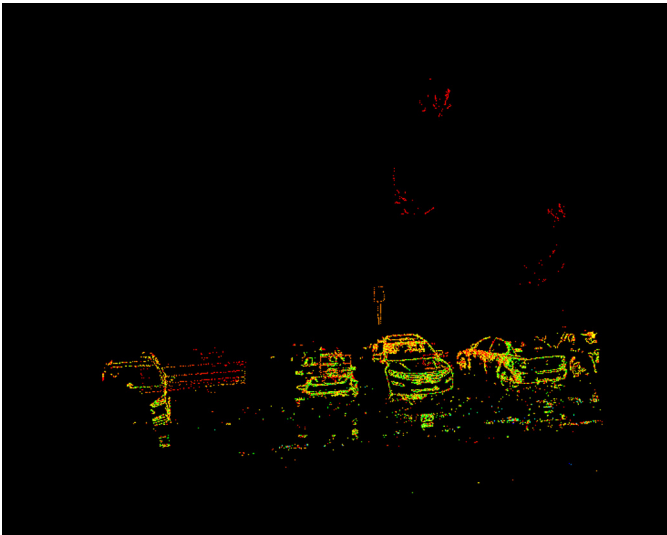


Fig. 29 Scene with only points of class 5 (mainly cars) being displayed.

it. We observe that almost all the snow has been filtered out of the points cloud. In Fig. 29, we display only the points associated to class 5 and we observe that almost only the cars have been singled out of the points cloud.

The method is thus shown to work on real complex scene data degraded by snow. It could be seen as a quick first step into a segmentation/classification process, which could be pressed forward after selecting points of one or the other class.

## 5 Conclusion

We have integrated into the lidar technology the statistical models already much in use in the radar technology. We have presented the basics of the detection theory: FAR, CFAR, and ROC curves. We have given three examples of simulations illustrating the use of these statistical models either for designing lidars or for preparing lidar missions. In view of making the preparation for a mission as adequate and realistic as possible, a simulator has been developed over the last six years. We have seen in Sec. 4 that the analysis of the simulation data shows that the simple Gaussian models developed for the targets were somehow defective.

Improved rendering of the overall environment for the mission may be obtained by using more sophisticated simulators such as one from the company Cognata for ADAS or the Ondulus-Lidar simulator developed by the Canadian company Presagis for helicopter pilots. A thorough analysis of results out of simulators could help develop more accurate statistical models for the scenes to be visited prior to a mission.

We have also introduced the idea of developing mixtures of statistical models as an approach to thresholding and object classification at mission time. In this last topic, much remains to be done mainly for the accurate location of the peaks (clusters of occurrences) in the histogram, which may easily get more complex as was the case for the scene of Fig. 24 with snow.

## 6 Appendix A

In the case of a Gaussian PDF with zero mean, which is the PDF of the noise, the right-tail probability is the Pfa and is given by the equation:

$$\text{Pfa} = Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp^{-\frac{1}{2}t^2} dt. \quad (5)$$

For the  $H_1$  hypothesis, where the mean  $\mu$  is different from zero, the equivalent of  $Q(x)$  would be the Pd and would be given by the equation:

$$\text{Pd} = Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp^{-\frac{1}{2}(t-\mu)^2} dt. \quad (6)$$

The relation between the threshold  $\gamma$  and the Pfa for Gaussian PDFs is given as

$$\text{Pfa} = Q(\gamma) = \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}} \exp^{-\frac{1}{2}t^2} dt. \quad (7)$$

This equation can be inverted to decide the value of the threshold for a tolerable Pfa:

$$\gamma = Q^{-1}(\text{Pfa}), \quad (8)$$

where  $Q^{-1}(x)$  is the inverse function of  $Q(x)$  and is said to always exist for Gaussian PDFs [Ref. 18, p. 21]. These two functions are the error function  $\text{erfc}(x)$  and its inverse  $\text{iерfc}(x)$ . Hence, for the case of Gaussian PDFs, solutions are easily found.

If we call  $p(x; H_1)$  the probability that the value  $x$  belongs to the  $H_1$  hypothesis and  $p(x; H_0)$  the probability that the value  $x$  belongs to the  $H_0$  hypothesis instead, the Neyman–Pearson theorem claims [Ref. 18, p. 65] “To maximize Pd for a given Pfa, choose  $H_1$  if

$$\frac{p(x; H_1)}{p(x; H_0)} > \gamma, \quad (9)$$

where the threshold  $\gamma$  is set as per Eq. (8).

## References

1. "Recommended Solutions to Address TAWS Manual Inhibition CFIT Cases Raised by NTSB Safety Recommendations A-17-035 and A-18-015 and GAJSC CFIT WG SE-54," RTCA Paper Number: 129-20/PMC-2018.
2. T. Nguyen et al., "A review of situation awareness assessment approaches in aviation environments," *IEEE Syst. J.* **13**, 3590–3603 (2019).
3. D. K. Barton, *Modern Radar System Analysis*, pp. 1–590, Artech House Inc (1988).
4. M. Skolnik, *Radar Handbook*, 2nd ed., pp. 2.1–2.68, McGraw Hill (1990).
5. X. Zhu, P. Church, and M. Labrie, "LIDAR for obstacle detection during helicopter landing," *Proc. SPIE* **6950**, 69500T (2008).
6. M. A. Vaughan et al., "Fully automated detection of cloud and aerosol layers in the CALIPSO lidar measurements," *J. Atmos. Ocean. Technol.* **26**, 2034–2050 (2009).
7. C. Goodin et al., "Predicting the influence of rain on LIDAR in ADAS," *Electronics* **8**, 89 (2019).
8. W. Song et al., "The effect of fog on the probability density distribution of the ranging data of imaging laser radar," *AIP Adv.* **8**, 025022 (2018).
9. B. Blagojevic, "Modeling of fluorescence lidar ROC curves," *IEEE Sens. J.* **10**(3), 772–778 (2010).
10. F. Zhang et al., "Anti-dynamic-crosstalk method for single photon LIDAR detection," *Proc. SPIE* **10605**, 1060503 (2017).
11. J. D. Allen et al., "A compressed sensing method with analytical results for lidar feature classification," *Proc. SPIE* **8055**, 80550G (2011).
12. K. K. Hon et al., "Application of short-range LIDAR in early alerting for low-level windshear and turbulence at Hong Kong International Airport," *Hindawi J.* **2014**, 162748 (2014).
13. G. Roy et al., "Physical model of snow precipitation interaction with a 3D lidar scanner," *Appl. Opt.* **59**(25), 7660–7669 (2020).
14. L. A. Klein, *Sensor and Data Fusion for Intelligent Transportation Systems*, pp. 1–235, SPIE (2019).
15. G. Melotti, A. Asvadi, and C. Premevida, "CNN-LIDAR pedestrian classification: combining range and reflectance data," in *IEEE Int. Conf. Veh. Electron. and Saf.* (2018).
16. R. D. Richmond and S. C. Cain, *Direct-Detection LADAR Systems*, pp. 1–138, SPIE Press (2010).
17. S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I, Estimation Theory*, pp. 1–595, Prentice Hall PTR (1993).
18. S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II, Detection Theory*, pp. 1–560, Prentice Hall PTR (1998).
19. G. Roy et al., "Enhanced scanning agility using a double pair of Risley prisms," *Appl. Opt.* **54**(34), 10213–10226 (2015).
20. W. Wagner et al., "Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner," *ISPRS J. Photogramm. Remote Sens.* **60**, 100–112 (2006).
21. Q. Li, "Decomposition of airborne laser scanning waveform data based on EM algorithm," *Int. Arch. Photogramm. Remote Sens. and Spatial Inf. Sci.* **XXXVII**(Part B1), pp. 211–217 (2008).
22. G. Tremblay, X. Cao, and R. Bernier, "3D resolution in scanning lidars - experiment and MC simulations (U)," DRDC-RDDC-2016-C314, (U) (limited distribution) (2016).
23. J. J. Oliver, R. A. Baxter, and C. S. Wallace, "Unsupervised learning using MML," in *Proc. 13th Int. Conf. Mach. Learn.* (1996).

**Robert Bernier** received a master's degree in physics from Laval University, Québec City, Canada. He has accumulated 34 years of experience in optical interferometry, laser-ultrasonics, lidar, machine vision, millimeter wave radar including unmanned vehicle radar and high-level R&D contracts management. He has also for 15 years taught physics at the pre-university level.

**Xiaoying Cao** received a PhD in chemical engineering from Queen's University, Kingston, Ontario, Canada. She has been working in lidar data analysis and modeling in past 20 years. She is now working as a lidar consultant in Ontario, Canada. Her current research interests are measurements and study of scattering properties of aerosols, atmospheric dispersion modeling, lidar simulation and application in aerosol and target detection.

**Gilles Roy** received a PhD in physics from Laval University, Québec City, Canada. He has accumulated over 40 years of experience in lidar technology, aerosols scattering, multiple scattering and depolarization.

**Grégoire Tremblay** received his PhD from Laval University, Québec City, Canada, in 1999 for his work in astronomical instrumentation. He worked for four years at EXFO as senior optical designer. Afterward, he worked for 15 years as a senior physicist at Aerex Avionics. He now works at Defence Research and Development Canada, Valcartier Research Centre, where his main interests are lidar, aerosol scattering, and atmospheric propagation.