WILEY | Hindawi

Review Article

Sizing of a Standalone PV System with Battery Storage for a Dairy: A Case Study from Chile

Pablo Viveros ,[1] **Francisco Wulff** ,[1] **Fredy Kristjanpoller** ,[1] **Christopher Nikulin** ,[2] **and Tomás Grubessich** [1]

[1]*Department of Industrial Engineering, Universidad Técnica Federico Santa María, Avenida España, Valparaíso 1680, Chile*
[2]*Product Design Engineering Department, Universidad Técnica Federico Santa María, Avenida España, Valparaíso 1680, Chile*

Correspondence should be addressed to Pablo Viveros; pablo.viveros@usm.cl

In this paper, a stochastic simulation model for a standalone PV system sizing is replicated and extended to supply a dairy's power demand. A detailed hourly-based simulation is conducted considering an hourly load profile and global solar radiation prediction model. The stochastic simulation model is based on a thorough statistical analysis of the solar radiation data and simulates the energy yield, the excess energy curtailed, and the state of charge of the batteries for the sizing month and the whole year, providing the designer autonomy factor values d to properly size the PV system, finding the optimum combination of installed peak power (P_m) and battery storage capacity (C_L) that meets the application load requirements, considering a preset reliability level at minimum cost. The model makes use of the NASA'S Surface Meteorology and Solar Energy database to obtain solar radiation data. Results show a substantial reduction of 44% in installed peak power and battery storage capacity when compared to conventional methodologies, considering three days of autonomy, and an 85% reduction considering four days. Considering the goodness of fit test results, the Wakeby distribution best represents the behavior of historical solar radiation data for the site in almost half of the months. This article seeks to contribute to the literature gap in the application of methodologies for the multicomponent power supply in the dairy industry through the use of renewable energy.

1. Introduction

Over the last decades, the amount of energy consumed by different types of industries has been growing enormously due to the economic growth of different countries, which in turn is triggered by the increase of world population and consumption patterns [1]. The environmental impact of human activities is contributing to rapid climate change due to CO_2 emissions, released as a result of fossil fuel combustion [2]. Solar energy is freely available and environmentally friendly, being widely adopted as an alternative for conventional electricity generation.

The sizing of standalone PV systems is an important task for the designers to find the most reliable and profitable combination of peak power and battery storage capacity that properly meets load demand, avoiding an equipment failure occurrence, considering the random behavior of solar

radiation [3]. Therefore, the objective is to improve energy efficiency in terms of economic performance and reliability standards to ensure system autonomy.

In this sense, several methods have been proposed and developed in the literature to size standalone PV systems, which differ according to the approach used.

Before PV system sizing, in the design step, it is essential to identify the inherent losses present in the whole system to adequately size the system components. As for PV panels, the input energy may be reduced by shading, reflection due to the angle of incidence, dust, degradation (ageing), temperature effects, etc. Before the converted energy reaches the load demand, power losses due to the efficiency of the power conditioning units (inverters and charge controllers) need to be considered [4]. As for batteries, the charge/discharge process may lead to some power losses due to its internal resistance [5] and others in the battery storage system due to

the natural degradation of battery capacity [5]. Since the above thing directly affects PV system performance, the task consists of controlling the system operation to minimize output power losses.

It is also important to take into account the balance between energy yield by the PV array and energy consumed by the load. Any short-term mismatch between them must be counterbalanced usually by rechargeable batteries. It is for this reason that a factor of autonomy days or days of battery backup, d, has been proposed in the literature [5] to represent the number of days on which the PV system can meet the load demand in the absence of power generation and Peak Sun Hours (PSH). In other words, d represents the maximum number of days that the PV system can supply energy to loads independently.

In this sense, the research carried out by Kaplani and Kaplanis [6] presents a stochastic simulation model to determine the size of autonomous photovoltaic systems in terms of peak power (P_m) and battery storage capacity (C_L) for the critical month, thus ensuring compliance with annual reliability standards. The model, in addition to considering the days of autonomy as an exogenous variable, we consider as main inputs the analysis and simulation of daily-based solar radiation and the load profile, obtaining, as a result, the state of charge of the battery, the excess energy burnt, the energy losses, and the P_m and C_L results for sizing the system.

The main contribution of this article lies in the lack of the literature that focuses exclusively on standalone photovoltaic system sizing methods and their implementation in the dairy industry, using direct primary energy to satisfy the simultaneous demand of multiple components for the milking process [7]. Although research such as [8, 9] addresses the sizing of PV systems in dairies, they only focus on supplying certain specific components of the milking process, unlike this research. On the other hand, this article seeks to extend the original method of [6], incorporating an hourly load profile, and extending the stochastic simulation elaborated in this article to consider an hourly basis for the determination of the capacity indicators. In order to meet these goals, there were three primary objectives of this study:

(1) To size a standalone PV system for a dairy in the south of Chile by replicating the method proposed by [6], in conjunction with reliable hourly solar radiation prediction data, performing a detailed hourly-based simulation and deepening, thus, the model proposed by these authors

(2) To test the hourly-based proposed model for a different geographical location than the original study to supplement and extend its findings

(3) To validate if the results obtained from the hourly-based method reach similar reductions in installed peak power and battery storage capacity as compared to conventional methodologies, as demonstrated by the daily-based method proposed by [6]

The methodology is based on stochastic modeling of the performance of both of the PV array and the battery storage system and provides the energy delivered, the excess energy curtailed, the load profiles, and the state of charge of the batteries for the sizing month and the whole year, as well as data on the success rate for the determination of the autonomy factor d and the sizing of the system. To generate hourly solar radiation data, the model proposed in [10] is employed.

2. Literature Review

2.1. Standalone Photovoltaic Systems. Photovoltaic (PV) power generation systems are currently one of the systems with the fastest growth in the use of direct solar energy, considering the latter a clean, environmentally friendly source of energy, and abundantly available in most locations in the world [11]. These systems are fundamentally designed to transform solar radiation useful energy into electricity, through the use of photovoltaic panels. In addition to the panels, PV systems are composed of a current controller, an inverter to supply direct and alternating current loads, and a storage battery.

Although PV systems generally require a high initial investment, they can offer a cost-effective alternative to certain isolated off-grid locations, in comparison to the installation of power lines with central distribution [11]. In such scenarios, standalone photovoltaic (SAPV) systems become a plausible alternative since they can supply off-grid power generation. Therefore, the reliability of these systems becomes a relevant aspect to address, understanding it as the system capacity to satisfy load demand during a certain evaluation period [12].

The sizing of photovoltaic systems seeks to determine the optimal combination of variables associated with the arrangement of photovoltaic panels, battery storage capacity, and tilt angle, among other parameters related to the components, to obtain the best compromise between power reliability and system cost. Given the vast diversity of approaches, Khatib et al. [11] classify the different sizing methods reviewed in the literature into intuitive methods, numerical methods, analytical methods, computational tools, computational intelligence methods, and hybrid methods. Likewise, the evaluation criteria are classified into technical parameters, considering the loss of load probability, loss of power supply probability, the state of charge (SOC), among others; economic parameters, considering the net present value (NPV) and the annualized cost; and social parameters, such as portfolio risk or even social acceptance.

Regarding the restrictions related to SAPV, the literature review developed by Fouad et al. [13] reveals that the factors with the greatest impact on the performance of photovoltaic systems analyzed are related to the battery efficiency and physical photovoltaic panels structure. Meanwhile, Khatib et al. [11] add physical space budget and energy constraints related with load profile fulfillment in the case of SAPV as essential aspects considered in the reviewed literature.

2.2. PV Systems on Daily Farms. The use of PV systems can offer more than just a cost-efficient alternative versus conventional electricity generation systems. Recently, these

types of alternatives also respond effectively to a requirement promoted both by the scientific community and by international organizations for the mitigation of anthropogenic emissions. One of the strategies recommended in research such as [14] is the proposal of primary energy saving measures through the implementation of PV systems for power generation and consumption in dairy farms.

In this sense, research such as developed by Breen et al. ([15], cited in [7]) reveals the potential of photovoltaic systems to substantially reduce carbon dioxide emissions in dairy farms, one of the main current anthropogenic emissions. In the particular case of this industry, a growth of 22% in milk production is expected for the 2018–2027 period [16], and 19% in its consumption by 2050 ([17], cited in [7]). From the energy consumption perspective, it is added the fact that electrical energy represents one of the items with the highest consumption for both conventional and organic dairy farms, representing 14% and 24% of total primary energy consumption, respectively, according to [18]. Considering the above, the incentives for photovoltaic system installation, particularly in the dairy industry, are considerable.

Regarding the review of related research, case studies such as De Blas et al. [8] and Zhang et al. [19] address the implementation and sizing of photovoltaic systems supplying certain components or subsystems within diary productive system, namely, the pumping system and the cooling system, respectively. Breen et al. [7], meanwhile, elaborate a generalizable and scalable optimization model for PV system sizing in dairy farms, using for the definition of the objective function the weighted sum method, considering both a financial criterion and the contribution ratio in terms of autoconsumed energy with respect to gross annual electricity consumption, using a trade-off parameter to search for and analyze Pareto-optimal solutions. Unlike the previously reviewed models, the model implemented in this research considers an autonomous system, which is why the need to ensure reliability and autonomy arises before the consideration of financial evaluation criteria or the contribution from renewable energy.

2.3. Estimation of Global Solar Radiation.

Solar radiation measuring results are essential for the design and operation of solar energy technologies, especially for the implementation of solar collectors and the sizing of photovoltaic systems. However, the measurement of this magnitude, despite being accurate, is not always available for modeling and forecasting, mainly due to the high initial investment and maintenance costs for its measurement and recording instruments [20]. For this reason, the alternative that arises within the scientific community is to correlate radiation with other available meteorological parameters, developing a wide variety of models based on this approach.

One of the first empirical models is the one developed by Angstrom [21] for the measurement of monthly average daily global radiation, developing an empirical correlation model considering the effect of the daily average of clear-sky radiation and the sunlight duration fraction. Meanwhile,

Page [9] decides to modify the correlation model, replacing the clear-sky radiation with the daily average extraterrestrial radiation on a horizontal surface. Since then, a series of empirical models have been developed that have modified and extended the Angstrom–Page model, including new parameters or different mathematical expressions to establish empirical relationships. See, for example, [22, 23] reviews that demonstrate the great variety of empirical models carried out to date, as well as the location dependence of the parameters.

Another alternative that has emerged in recent decades for estimating and forecasting global radiation is the use of artificial neural networks, where the inputs usually correspond to a great variety of meteorological and geographical parameters, while the output corresponds to the global solar radiation for different time scales [24]. The main advantages of these models for estimating global radiation are their location adaptability, and a perceptible accuracy improvement compared to traditional empirical models. However, the disadvantage of these models is its complexity and the high number of parameters in order to achieve such accuracy, not clearly establishing the effect of this inclusion, in comparison to empirical models [20].

In an attempt to classify for the first time the vast variety of models for global radiation in terms of the time span contemplated and the type of method approached, Zhang et al. [20] develop a literature review classifying them according to the output into monthly average of daily global radiation, daily global radiation, and hourly global radiation method. The authors also consider two main categories regarding the type of estimation: empirical models, which are disaggregated into sunlight duration fraction (SDF) models, modified models (MSDF), nonsunshine duration fraction (NSDF) models, and artificial neural networks (ANNs). The study also presents performance indicators such as the root mean squared error (RMSE) and the mean absolute percentage error (MAPE) to determine the accuracy of each reviewed model based on observed and predicted values. The research concludes that the classical models of fraction of solar duration (SDF), as well as artificial neural networks (ANNs), have the best performance in terms of RMSE and MAPE. Although SDF models have slightly lower performance than ANN models, the latter generally requires a large amount of data for the training phase to improve the performance indicators presented.

Considering this last classification, the model developed by Kaplani and Kaplanis [6] falls into the category of the long-run monthly average daily global radiation. The methodology obtains a representative value of global daily radiation throughout a month, then averaging the values obtained year by year, for each month. The summarized methodology associated with this model, as well as the nomenclature used and replicated in this investigation, is presented in Section 6 and Appendix A, respectively.

The aforementioned sizing method considers a monthly fitting for daily solar radiation using probability distributions. In this regard, specialized distributions such as the Wakeby distribution have proven their effectiveness in particular for power generation modeling in planning

applications of PV power systems [25]. Also, the Johnson SB and Generalized Extreme Value distributions are usually considered for modeling meteorological measurements such as wind speed, generally providing a superior fit to one-component probability density functions [26].

3. Design of the Dairy

The case study is carried out in a dairy located in the city of Río Bueno, in the south of Chile, with coordinates south latitude 40.331 and east longitude 72.498. In the economic field, agricultural equipment prices have substantially increased and, conversely, milk prices have continued to decrease considering the producer's perspective. There are few cost-saving opportunities in these industries, being necessary to analyze more cost-effective options for power generation. In this respect, Chile is endowed with abundant solar radiation, and the amount of electrical energy that can be generated from solar panels is potentially significant.

The electrical load of the diary consists of equipment for milking, cooling, cleaning, water heating, illuminating, and ventilation, among others, such as milking machine, vacuum and water pump, cooling tank, cleaning systems, and light bulbs. The dairy performs two milkings per day, from 5:30 to 7:15 and from 15:30 to 17:15 considering ordinal hours. The milking process is described below, as well as the energy consumption required by the dairy equipment for this case study:

(i) Milking machine: its main function is the aspiration to activate the teat cups, directing the milk obtained towards the cold tank for subsequent accumulation. The machine has 14 units and is equipped with a vacuum pump, which has a 2.2 kW electric motor including an additional 10% for the operation of the pulsators. Also, a 0.5 HP motor is added to drive the milk to the cold tank. Its operation time matches the milking process, from 5:30 to 7:15 and from 15:30 to 17:15.

(ii) Cold tank: it is the device that allows the storage of milk in optimal conditions. It is equipped with a 4 kW engine and has a capacity of 4,275 liters. It has two 70 W agitators each and a ventilation system, which consists of two 200 W fans each. The operation time takes place after the milking, and all its components operate simultaneously. The period is between 7:15 and 11:00 and between 17:45 and 22:00.

(iii) Lighting: the lighting includes the use of luminaires for the milking room, a corridor, and the machine room that contain the devices described above. This item gives a total of 472 W, but it is estimated that it reaches around 600 W by adding the warehouse and surroundings. The lighting schedule coincides with the milking in the morning, that is, between 5:30 and 7:15 hrs.

(iv) Hot water: the warm water is used for machine cleaning and to wash the cow's udders. The water heater demands 1500 W power supply, and it is estimated that it works approximately 6.4 hours per day, to maintain the water temperature at about 65°C. Schedules in which the water heater operates are between 06:45 and 08:25 and between 16:45 and 18:25 hrs (milking process), and between 13:00 and15:45 hrs (cold tank).

(v) Water pump: in order to supply the dairy with water, a 1 HP submersible pump is used. It exceeds a total height of 25 meters, with a flow of $3 \, m^3$/hr. Average daily water consumption corresponds to $11 \, m^3$. Therefore, the pump operates for 3 hours and 40 minutes, consuming approximately 2.74 kW per day. The hours are usually between 06: 45 and 09:15 and between 16:45 and 17:55.

(vi) Machine washing: to keep the tank in optimal conditions, it must be washed after milk collection. The closed tanks have an automatic washing system, which uses hot and cold water combined with acidic and basic cleaning products. The motor that allows this is 1 HP. The duration of the tank washing is 45 minutes, immediately after the collection by the collection truck, which occurs around 12:30 pm. In the case of the milking machine, the washing seeks to remove organic matter, fats, and minerals for the milk. The purpose is to prevent the proliferation of bacteria. The automatic washing is carried out, using a 0.5 HP automatic machine minutes after each milking. The period of milking machine washing is 30 minutes, at 08:00 hrs and 18:00 hrs.

Table 1 summarizes the consumption and operating times associated with dairy equipment.

The total load demand for the dairy was estimated to be 59.83 kWh/day, which corresponds to the peak summer demand. Summer has the highest energy consumption compared to other seasons. This is because the cooling tank has to operate for a longer period to keep milk at an appropriate temperature. Since the dairy operates at certain times of the day, especially during milking, the building of a detailed hourly load profile is required and presented in Figure 1.

4. Solar Radiation

4.1. Solar Radiation Data Analysis. In this study, based on the model proposed by [6], the daily global solar radiation data (H) were obtained from the NASA's Surface Meteorology and Solar Energy database [27] for the dairy located at south latitude 40.331 and west longitude 72.948, considering the period between 1985 and 2004. For each of the 20 years, based on the recommended average days for months [28] as cited in [29], the average day for each month, including the day both before and after the average day (i.e., three consecutive days), was taken to form the global solar radiation data for the analysis. The global solar radiation H for each month at the site concerned is shown in Figure 2.

For each month, solar radiation data were fitted by MLE (see Appendix C) considering six continuous probability

TABLE 1: Consumption and operation times of the diary's equipment.

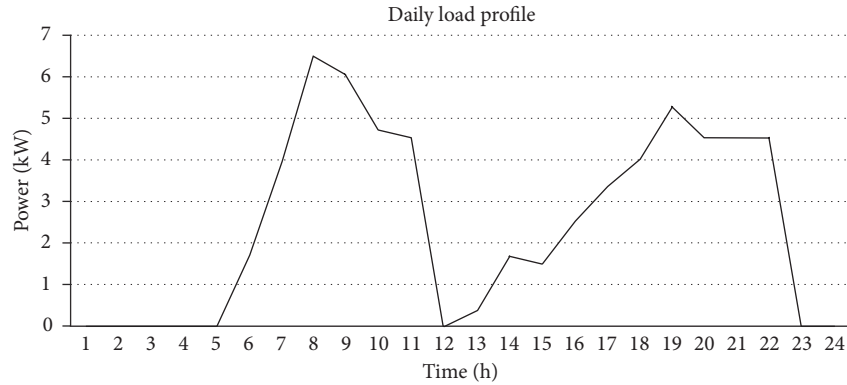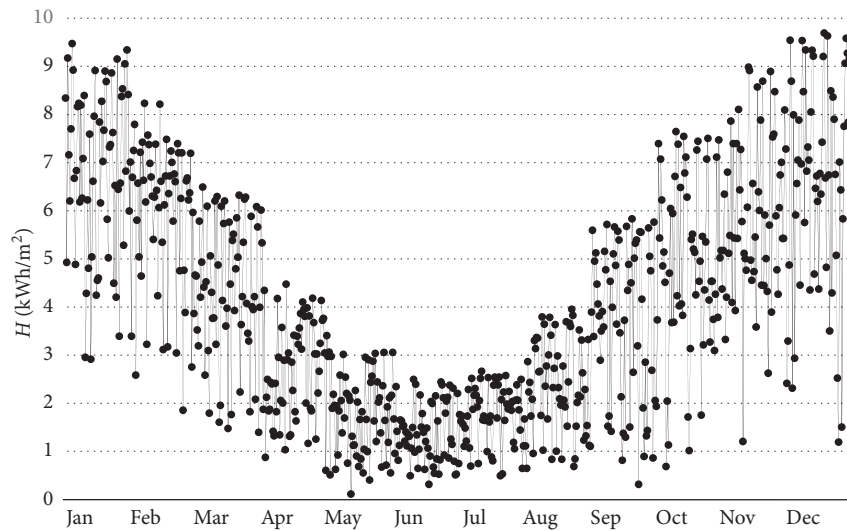| Equipment	Consumption (kW)	Operation times
Cold tank	4.54	07:15–11:00; 17:45–22:00
Milking Machine	2.573	05:30–07:15; 15:30–17:15
Hot water	1.5	06:45–08:23; 16:45–18:23 (milking) 13:00–15:45 (tank)
Lighting	0.6	05:30–07:15
Water pump	0.746	06:45–9:15; 16:45–17:55
Others (Machine washing)	0.373–0.746	08:00–08:30; 18:00–18:30 (milking) 12:30–13:15 (tank)

FIGURE 1: Load profile of the dairy.

FIGURE 2: Daily global solar radiation H on an average day, including the day both before and after the daily average per month for the years 1985–2004 for the site concerned.

distributions, namely, Beta, Generalized Extreme Value, Johnson SB, Normal, Wakeby, and Weibull distributions (see Table 2), and were tested according to the Kolmogorov–Smirnov test (K-S test). The K-S statistic values are shown in Table 3. Then, the p value is obtained and used as selection criteria. It is important to notice that in this case, the null hypothesis H_0 establishes that the data follow the specified distribution for each distribution; therefore, the selection must consider a p value that cannot reject that hypothesis (see Appendix D). Following this observation,

the distribution with the highest possible p value is chosen for each month.

The p value was calculated using a novel method named Exact-KS-FFT, expressing the p value as a double-boundary noncrossing probability for a homogeneous Poisson process, which is then efficiently computed using fast Fourier transform [30]. The method has proved his accuracy and numerical efficiency versus the algorithm proposed by Marsaglia et al. [31]. The results for the fitted distributions are presented in Table 4.

TABLE 2: Probability distributions to average daily global solar radiation modeling, for each month.

Distribution	Probability density function	Parameters
Beta	$f(x; a, b, \alpha_1, \alpha_2) = (1/(B(\alpha_1, \alpha_2))) \cdot ((x-a)^{\alpha_1-1}(b-x)^{\alpha_2-1}/(b-a)^{\alpha_1+\alpha_2-1})$, $B(\alpha_1, \alpha_2) = \int_0^1 t^{\alpha_1-1}(1-t)^{\alpha_2-1} dt$;	a, b: distribution Limits $\alpha_1, \alpha_2 > 0$
Generalized Extreme Value	$f(x, k, \mu, \sigma) = \begin{cases} (1/\sigma)\exp(-(1+kz)^{-(1/k)}(1+kz)^{-1-(1/k)}), & k \neq 0 \\ (1/\sigma)\exp(-z - \exp(-z)), & k = 0 \end{cases}$	$z = ((x-\mu)/\sigma)$ k, μ, σ
Johnson SB	$f(x; y, \delta, \gamma, \xi) = (\sigma/(\lambda\sqrt{2\pi} z(1-z)))\exp(-(1/2)(\gamma + \delta \ln(z/(1-z)))^2)$	$z = ((x-\xi)/\lambda)$ y, δ, γ, ξ
Normal	$f(x; \mu, \sigma) = (1/(\sigma\sqrt{2\pi}))e^{((-(x-\mu)^2)/(2\sigma^2))}$	μ, σ
Wakeby	$f(F; \alpha, \beta, \gamma, \delta, \xi) = \xi + (\alpha/\beta)(1 - (1-F)^\beta) - (\gamma/\delta)(1 - (1-F)^{-\delta})$	$\alpha, \beta, \gamma, \delta, \xi$
Weibull	$f(x; \alpha, \beta) = (\alpha/\beta)(x/\beta)^{\alpha-1}\exp(-(x/\beta)^\alpha)$	α, β

TABLE 3: The K-S statistic values of the fitted distributions on the pdf of each month for the site.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Beta	0.115	0.114	0.093	0.054	0.147	0.079	0.122	0.057	0.052	0.126	0.099	0.079
Gen. Extreme Value	0.073	0.080	0.076	0.074	0.067	0.082	0.061	0.091	0.093	0.086	0.064	0.080
Johnson SB	0.055	0.106	0.072	0.062	0.076	0.055	0.081	0.055	0.050	0.094	0.075	0.085
Normal	0.096	0.175	0.097	0.083	0.071	0.089	0.105	0.107	0.124	0.084	0.072	0.103
Wakeby	0.053	0.107	0.079	0.062	0.069	N/A	0.075	0.054	0.048	0.078	0.081	0.068
Weibull	0.082	0.166	0.111	0.100	0.126	0.087	0.136	0.086	0.172	0.144	0.077	0.166

TABLE 4: p values associated to the K-S statistic value for the fitted distributions.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Beta	0.377	0.387	0.643	0.991	0.135	0.820	0.308	0.984	0.994	0.273	0.565	0.820
Gen. Extreme Value	0.883	0.808	0.853	0.873	0.934	0.784	0.969	0.669	0.643	0.734	0.953	0.808
Johnson SB	0.989	0.478	0.893	0.964	0.853	0.989	0.796	0.989	0.997	0.630	0.863	0.747
Normal	0.604	0.045	0.591	0.772	0.902	0.695	0.490	0.466	0.290	0.759	0.893	0.515
Wakeby	0.993	0.466	0.820	0.964	0.919	N/A	0.863	0.991	0.998	0.831	0.796	0.926
Weibull	0.784	0.065	0.420	0.552	0.273	0.721	0.198	0.734	0.051	0.150	0.842	0.065

For the months of January, August, September, October, and December, the Wakeby distribution provides the best fit, while for the months of February, May, July, and November, the Gen. Extreme Value distribution works best. For March and June, the Johnson SB distribution is the most appropriate. Finally, for April the Beta distribution provides the best fit. Although the Normal distribution is often chosen to fit the data, in this case, it does not present substantial differences in terms of the statistic for the rest of the fitted distributions. It has to be emphasized that while this study considers only these six types of distributions, other alternatives may eventually improve the fit.

As already stated, to achieve yearly performance, PV system sizing is carried out based on the "worst month" method, i.e., the month with the lowest average solar radiation levels. Thus, June is selected for sizing in the following sections.

4.2. Model for Estimating the Hourly Global Solar Radiation.

When hour-by-hour performance calculations need to be done, it may be necessary to start with daily data and then obtain hourly values from daily solar radiation estimation. The ratio of hourly global solar radiation in a day n_j, $I(h; n_j)$, to daily global solar radiation, $H(n_j)$, is calculated from equation (1) [29]. Baig et al. [10] modified the model proposed by Jain et al. [32] which tries to fit solar radiation to a Gaussian-type function, to better fit the recorded data during the start and the end periods of a day, as cited in [33]. In this model, r_h is estimated by equation (2), where h is the solar time, σ_G is the standard deviation of the Gaussian curve, and S_o is the day length of the day n_j, at a site with latitude ϕ, given by equation (3), where δ is the sun's declination. S_o correlates with σ_G, equations (4) and (5), based on the second version of the new approach to Jain's model as cited in [33]. $I(h; n_j)$ is determined from equation (1), leading to equation (6). The equations used in this section are listed as follows:

$$r_h = \frac{I(h; n_j)}{H(n_j)}, \tag{1}$$

$$r_h = \frac{1}{2\sigma_G \sqrt{2\pi}} \left\{ \exp\left(-\frac{(h-12)^2}{2\sigma_G}\right) + \cos\left[180 \frac{(h-12)}{(S_o - 1)}\right] \right\}, \tag{2}$$

$$S_o = \frac{2}{15}\cos^{-1}(-\tan\phi \cdot \tan\delta), \tag{3}$$

$$S_o = 4.054 \cdot \sigma_G, \tag{4}$$

$$\sigma_G = 0.246 \cdot S_o, \tag{5}$$

$$I(h, n_j) = r_h \cdot H(n_j). \tag{6}$$

5. Proposed Standalone PV System

The schematic representation of the proposed standalone PV system is shown in Figure 3. The main components are PV generator, battery system, and power conditioning units (inverters and charge controllers).

6. Stochastic PV Sizing Simulation Model

The stochastic PV sizing methodology corresponds to the model developed by [6], but considering some modifications specifically in steps 3 and 6 since a detailed hourly-based simulation is carried out. It consists of the following steps:

(1) An essential first step is to obtain a reliable database that contains appropriate solar radiation and temperature data for the years under consideration. The application site is defined by entering latitude and longitude values. Based on a thorough statistical analysis of the global solar radiation data H, a probability density function (pdf) is generated for each month in such a way as to find the best fitting curve to the solar radiation data for the site. Then, the parameters of the fitted distribution are extracted for all months.

(2) The global solar radiation mean (H_m) and the standard deviation (σ_H) are calculated for each month. The one that exhibits the lowest average daily solar radiation levels is selected for sizing.

(3) Depending on the application type, estimation of the energy demand is carried out based on technical datasheets and expert judgment. Unlike [6], an hourly load profile is used for a more detailed hourly-based simulation.

(4) Estimation of the correction factor F, which refers to the energy lost along the path from the PV generator to the loads during day-to-day operations, and of the correction factor F' for the operation route battery-DC/AC inverter-loads, calculated accordingly to [6]. Estimation of R_m, based on the beam and diffuse components of the monthly radiation data according to the Collares-Pereira and Rabl model as cited in [29] and stated in [6]. To estimate F, C_{T_c} is calculated from equation (7), where T_C is the cell temperature and TC is the temperature coefficient, obtained from the manufacturer's specifications. T_C is calculated using equation (8) and can be estimated from the maximum average ambient temperature T_a, the irradiance G under standard test conditions (STCs), and the nominal operating cell temperature (NOCT) as cited in [29]. Combining these two equations leads to equation (9):
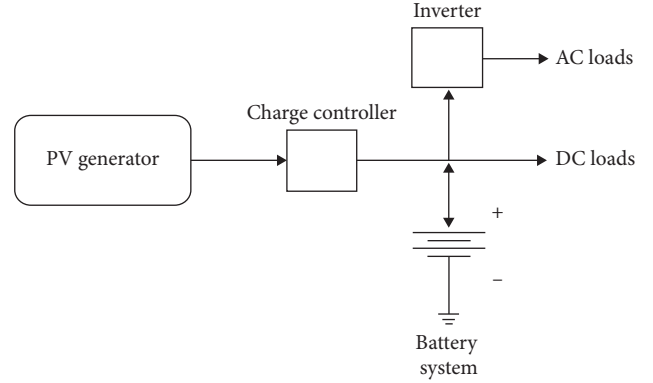

FIGURE 3: Standalone PV system configuration.

$$C_{T_c} = (T_C - 25°C) \cdot TC, \tag{7}$$

$$T_C = T_a + \left(\frac{NOCT - 20°C}{0.8(kW/m^2)} \right) \cdot G, \tag{8}$$

$$C_{T_c} = \left(\left(T_a + \left(\frac{NOCT - 20°C}{0.8(kW/m^2)} \right) \cdot G \right) - 25°C \right) \cdot TC. \tag{9}$$

(5) The lower and upper range values for the autonomy factor d are set depending on climatic conditions and the application in question. The higher the solar radiation, the lower the autonomy factor. d is set from d_{start} to d_{end}, ranging from 1 to 10. The range values are taken from [6].

(6) The simulation algorithm starts with iterating d from d_{start} to d_{end}. After each iteration, the corrected peak power $P_{m,cor}$ and the corrected battery storage capacity $C_{L,cor}$ are calculated with the H_m and σ_H previously calculated in step 2, as stated in [6]. Each iteration of d is comprised of 100 simulations for which the energy delivered E_{PV}, the hourly load demand $Q_{L,h}$, the energy loss E_{loss}, and the battery state of charge SOC are registered, for all hours of the month selected for sizing, to estimate the success rate in percentage provided by the specific value of d, for every hour of the month.

For each day of the month, the global solar radiation $H(n_j)$ is sampled from the selected distribution with parameters extracted in step 1. Therefore, for all hours of the month r_h is calculated using equation (2), thus obtaining hourly global radiation $I(h, n_j)$ for a particular day n_j. The hourly clearness index k_T is calculated using equation (10), where I_{ext} is the hourly extraterrestrial radiation on a horizontal surface for an hour between hour angles ω_1 and ω_2 [29]. For converting the hourly solar radiation from the horizontal surface to the inclined plane, tilted at

slope β from the horizontal, R is calculated using the isotropic diffuse model as derived by [34] from the isotropic model proposed by [35], as cited in [29], given by equation (11), where (I_b/I) is the fraction of the hourly radiation on a horizontal plane which is the direct beam, R_b is the geometric factor, i.e., the ratio of beam radiation on the tilted surface to that on a horizontal surface at any time, calculated based on the hour angle ω evaluated at the midpoint of the

hour, declination δ for the sampling day, site latitude ϕ, and slope β, I_d/I is the fraction of the hourly radiation on a horizontal plane which is diffuse, based on the Erbs et al. correlation [36] as cited in [10], given by equation (12), and ρ_g is the reflectance (the albedo) of the ground:

$$k_T = \frac{I(h, n_j)}{I_{ext}}, \tag{10}$$

$$R = \frac{I_b}{I}R_b + \frac{I_d}{I}\left(\frac{1 + \cos(\beta)}{2}\right) + \rho_g\left(\frac{1 - \cos(\beta)}{2}\right), \tag{11}$$

$$\frac{I_d}{I} = \begin{cases} 1.0 - 0.09k_T, & k_T \leq 0, \\ 0.9511 - 0.1604k_T + 4.388k_T^2 - 16.638k_T^3 + 12.336k_T^4, & 0.22 < k_T \leq 0.80, \\ 0.165, & k_T > 0. \end{cases} \tag{12}$$

For each hour, the energy delivered by the PV array E_{PV} is calculated using equation (13), and the remaining amount of energy after satisfying all load requirements is denoted by DE:

$$E_{PV} = P_{m,cor} \cdot I \cdot R. \tag{13}$$

To express the amount of energy remaining after consumption as a percentage, the auxiliary variable AUX is defined. In cases where the PV system is not generating any electricity ($E_{PV} = 0$), DE is estimated from equation (14), as the battery system is supplying the demand loads and AUX is estimated from equation (15):

$$DE = E_{PV} - Q_{L,h} \cdot F', \tag{14}$$

$$AUX = \frac{DE}{C_{L,cor} \cdot V}. \tag{15}$$

Alternatively, in cases where the PV system is generating power ($E_{PV} > 0$), DE is then estimated from equation (16). For the latter, three different cases are considered:

$$DE = E_{PV} - Q_{L,h} \cdot F. \tag{16}$$

(a) If DE < 0, then AUX is calculated using equation (17) since the battery bank must satisfy the load demand and the SOC of the battery is decreased:

$$AUX = \frac{DE \cdot F'}{C_{L,cor} \cdot V}. \tag{17}$$

(b) If DE = 0, then AUX is equal to zero since the energy yield by the PV system perfectly matches the load requirements.

(c) If DE > 0, then AUX is estimated using equation (15), since the PV system managed to cover all the loads and the SOC of the battery is increased.

Thus, the battery state of charge SOC is calculated from equation (18), considering the remaining energy, where SOC_a is the SOC at the moment immediately before the current period, that is, the previous hour, considering a simulation on an hourly basis. Any excess energy generated by the system is dissipated and is represented by E_{loss}, given by equation (19). For each simulation hour, the SOC is evaluated. If the value drops below the critical level 1-DOD, then a failure occurs and the simulation stops. The same process is repeated until the last hour of the month, recording the success rate for each value of d, from d_{start} to d_{end}:

$$SOC = SOC_a + AUX, \tag{18}$$

$$E_{loss} = (SOC - 1) \cdot C_{L,cor} \cdot V. \tag{19}$$

(7) Depending on the percentage of the time, it is desired to meet the load demand, and the minimum value of d that fulfills the system reliability criterion is selected.

(8) Based on the selected value of d obtained in step 7, the system sizing is carried out by substituting d into

$$P_{m,cor} = Q_L \cdot F \cdot \frac{(1 + \sqrt{d} \cdot 2 \cdot (\sigma_H/H_m))}{(PSH_m \cdot R_m)}, \qquad (20)$$

$$C_{L,cor} = Q_L \cdot F' \cdot \frac{(1 + \sqrt{d} \cdot 2 \cdot (\sigma_H/H_m))}{(V_{DC} \cdot DOD)}, \qquad (21)$$

calculating $P_{m,cor}$ and $C_{L,cor}$, as stated in [6].

(9) The simulation is performed for the remaining months of the year, taking into account the sizing parameters obtained in step 8. For each month, the goal is to obtain the success rate and the system's performance, that is, the E_{PV}, $Q_{L,h}$, E_{loss}, and SOC.

For each month, if simulation findings show that the system reliability criterion (success rate) for the selected value of d was not met, then return to step 7 and select the next incremented d.

These last two methodology steps correspond to a general vision of the system's performance during the whole year while ensuring the fulfillment of the criterion of "system reliability." The whole simulation model is presented in Appendix B. Please note that both flowchart and methodology correspond to the model proposed by [6], but with some modifications consistent with a detailed hourly-based simulation.

7. Simulation Results

Based on the "worst month" method, the month selected for sizing is June. The load demand was estimated to be 59.84 kWh/day. Assuming a maximum average ambient temperature in the daytime T_a of 8°C, calculated from the NASA's database for the years in question (1995 to 2004), as mentioned in the beginning of this study, the NOCT and the temperature coefficient TC are taken as (for example, from the technical specifications for the module PV-UD185MF5) 47.5°C and 0.452%/°C, respectively. Based on equation (9), C_{Tc} is calculated as 0.9215. $C_{charger}$, $C_{inverter}$, C_{bat-c}, C_{bat-d}, C_{pv-ag}, and C_{bat-ag} are taken as 0.98 [37], 0.9 [37], 0.95 [5], 0.95 [5], 0.9936 [38], and 0.8 [39], respectively. Thus, correction factors F and F' of 1.30 and 1.46, respectively, were calculated, based on [6]. It should be kept in mind that these values are customized to the application in question and are subject to the design, equipment, and technology used. The conversion factor R was calculated considering the photovoltaic panels inclined at an angle of 25° above the horizontal. This slope was calculated based on maximizing the total estimated energy for the winter months [29]. The batteries' discharge depth DOD was set to 80%, and the DC transfer voltage V was 120 V. The autonomy factor d was set from d_{start} to d_{end}, ranging from 1 to 10, which was increased by 0.5 each iteration. The success rates obtained for the corresponding iteration value of d with the $P_{m,cor}$ and $C_{L,cor}$ calculated using equations (20) and (21), respectively, are

shown in Figure 4. From this figure, the minimum value of d is selected to achieve at least 95% success for noncritical loads and at least 99% for critical loads. The success rates obtained for the $P_{m,cor}$ and $C_{L,cor}$, calculated by replacing d with its corresponding iteration value, are, respectively, depicted in Figures 5 and 6. It is worth highlighting that the PV system sizing, apart from being affected by the autonomy factor d, strongly depends on the weather conditions, since the $P_{m,cor}$ and $C_{L,cor}$ are also affected by the correction parameter (σ_H/H_m), that is, the global solar radiation data at the site. Regarding the first claim, it is logical that the more autonomy days a PV system has, the larger the size of the PV generator and battery bank will be.

For June, an autonomy factor d of 3 assures at least an average of 97% success for noncritical loads. The $P_{m,cor}$ calculated is 97.074 kWp and $C_{L,cor}$ is calculated to be 2.308 kAh. Contrasting the results with the conventional methodology [5], d_{cr} is calculated from linear equation from [6]. For noncritical loads, d_{n-cr} is estimated to be 4.532. Thus, P_m and C_L of 173.443 kWp and 4.124 kAh, respectively, are calculated, based on equations (1) and (2) from [6]. The proposed stochastic methodology, based on the model proposed by [6], far exceeds the conventional methodology, achieving a 44% reduction in installed peak power and battery storage capacity. Repeating the same analysis, for $d = 4$, the proposed methodology assures at least an average of 99% success for critical loads for June. In this case, the $P_{m,cor}$ and $C_{L,cor}$ obtained are 106.171 kWp and 2.525 kAh, respectively, as summarized in Table 5. From the conventional point of view, d is calculated using d_{cr} linear equation from [6], which results in $d_{cr} = 18.11$ and gives $P_m = 693.083$ kWp, based on equation (1) and $C_L = 16.481$ kAh, based on equation (2) from [6], as summarized in Table 6. Comparing the results, an 85% reduction in installed peak power and battery storage capacity is achieved with the proposed methodology, getting better results than with the conventional methodology.

The hourly performance of the proposed system throughout the sizing month June and $d = 3$, in terms of E_{PV}, $Q_{L,h}$, E_{loss}, and SOC, is shown in Figure 7. It can be seen from Figure 7 that the proposed standalone PV system successfully meets the load demand without the SOC falling below 60%. The battery bank was able to cover the energy demand for those hours at which no energy was generated by the PV modules. Besides, the amount of energy dissipated was considerable.

Conversely, Figure 8 shows a failed case where the low levels of solar radiation and, thus, the E_{PV} value caused the batteries to be unable to meet the demand in loads with a SOC falling below 20%.

7.1. Designed Standalone PV System Hourly Performance for the Other Months. The algorithm is run for the remaining months using the corresponding values of d, $P_{m,cor}$, and $C_{L,cor}$, as previously calculated, in such a way to achieve at least 95% and 99% of system autonomy. As the sizing is based on the month with the lowest solar radiation, it is expected to perform better. The success rates for each month
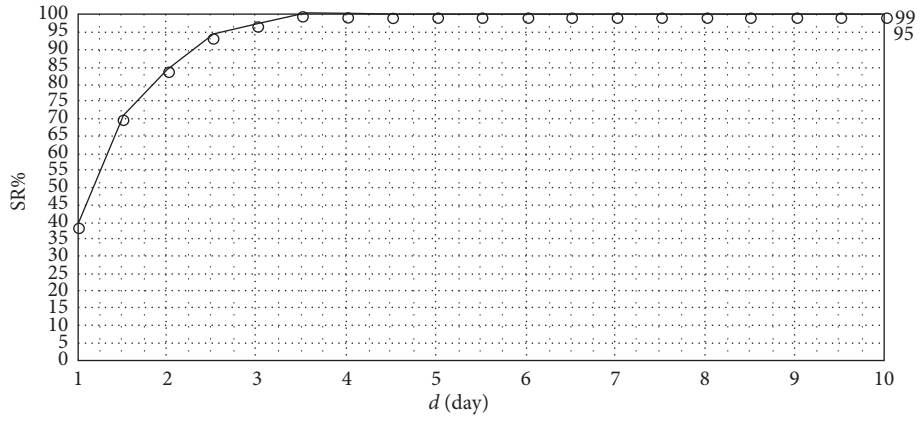
Figure 4: Success rates SR% vs different values for the autonomy factor d.
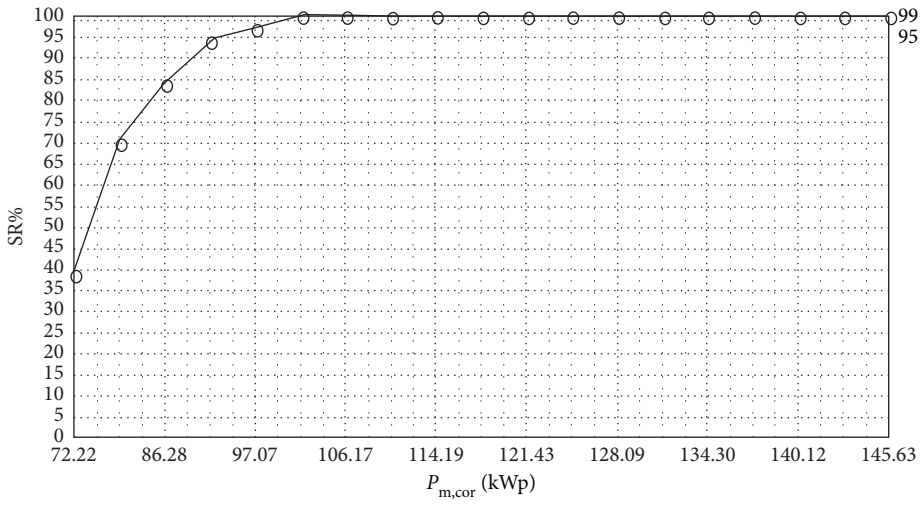

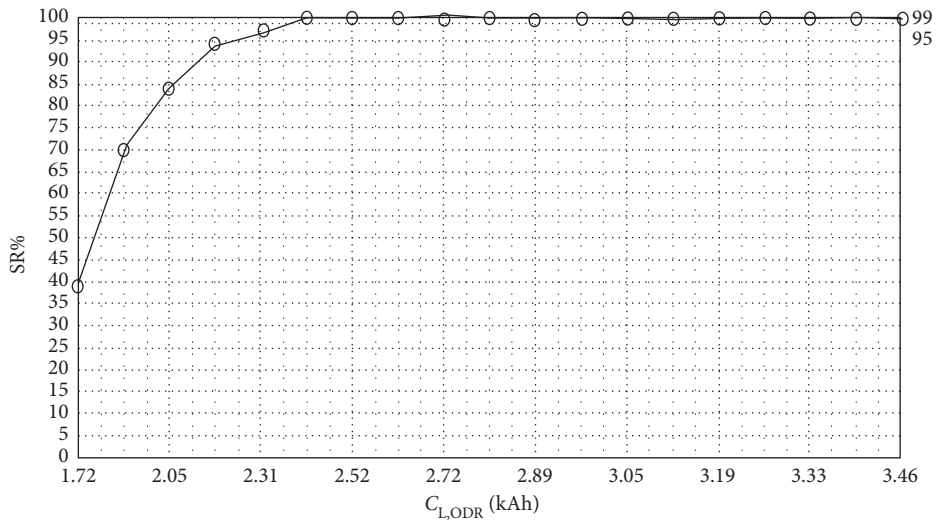Figure 5: Success rates SR% vs corrected peak power $P_{m,cor}$.

Figure 6: Success rates SR% vs corrected battery storage capacity $C_{L,cor}$.

TABLE 5: Success rates SR%, corrected peak power $P_{m,cor}$, and corrected battery storage capacity $C_{L,cor}$ values for different values of the autonomy factor d for the proposed stochastic methodology.

d	1	1, 5	2	2, 5	3	3, 5	4	4, 5	5
SR%	39	70	84	94	97	100	100	100	100
$P_{m,cor}$ (kWp)	72.22	79.85	86.28	91.95	97.07	101.79	106.17	110.29	114.2
$C_{L,cor}$ (kAh)	1.72	1.90	2.05	2.19	2.31	2.42	2.52	2.62	2.72

TABLE 6: Success rates SR%, peak power P_m, and battery storage capacity C_L values for different values of the autonomy factor d for the conventional methodology.

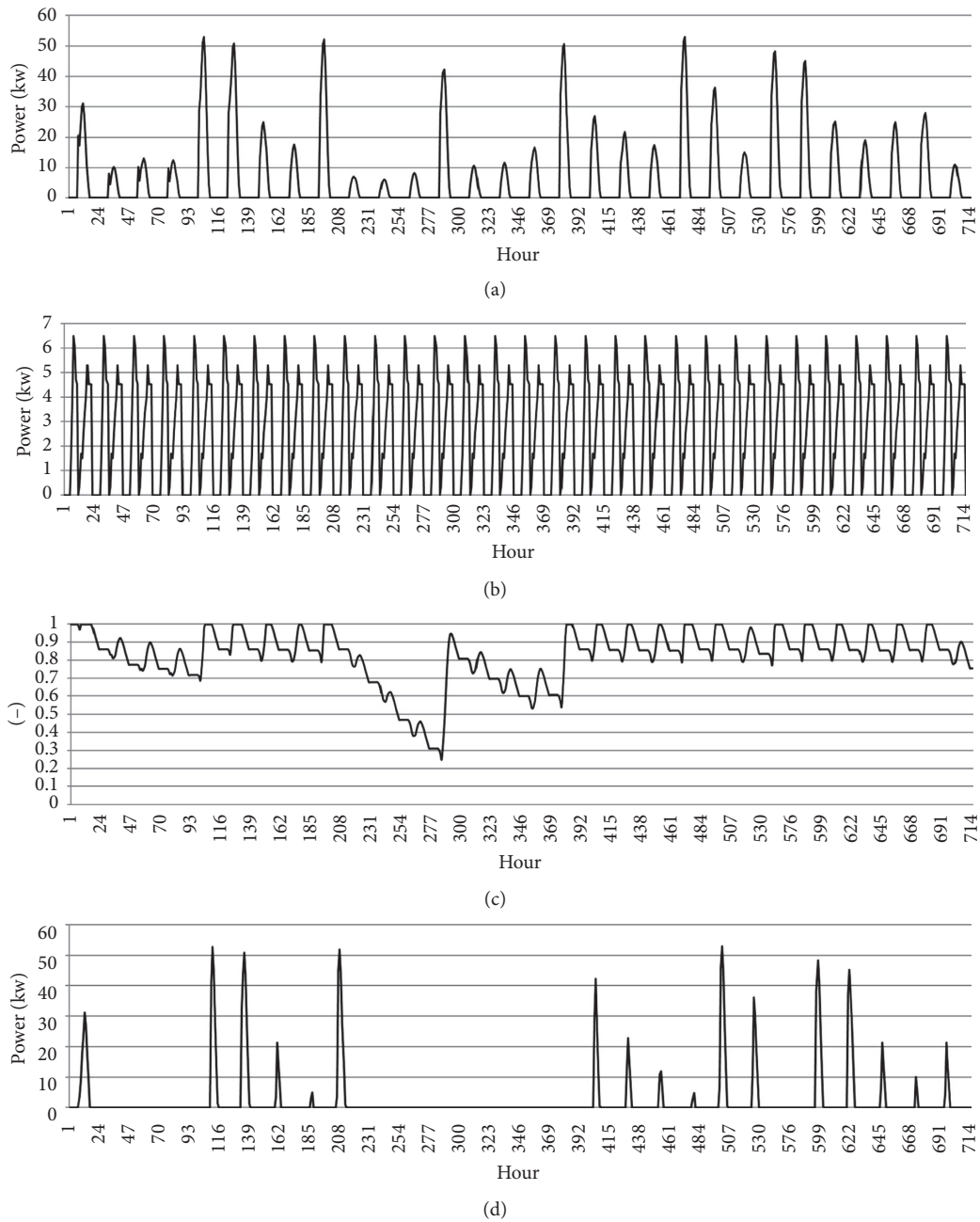| d	$d_{n-cr} = 4.53$	$d_{n-cr} = 18.11$
SR%	95	99
P_m (kWp)	173.44	693.08
C_L (kWA)	4.12	16.48

FIGURE 7: Designed system hourly performance for all hours of June, with $d = 3$, in terms of the energy delivered E_{PV}, load demand $Q_{L,h}$, excess energy curtailed E_{loss}, and state of charge of the batteries SOC. (a) E_{PV}, (b) Q_L, (c) SOC, and (d) E_{LOSS}.
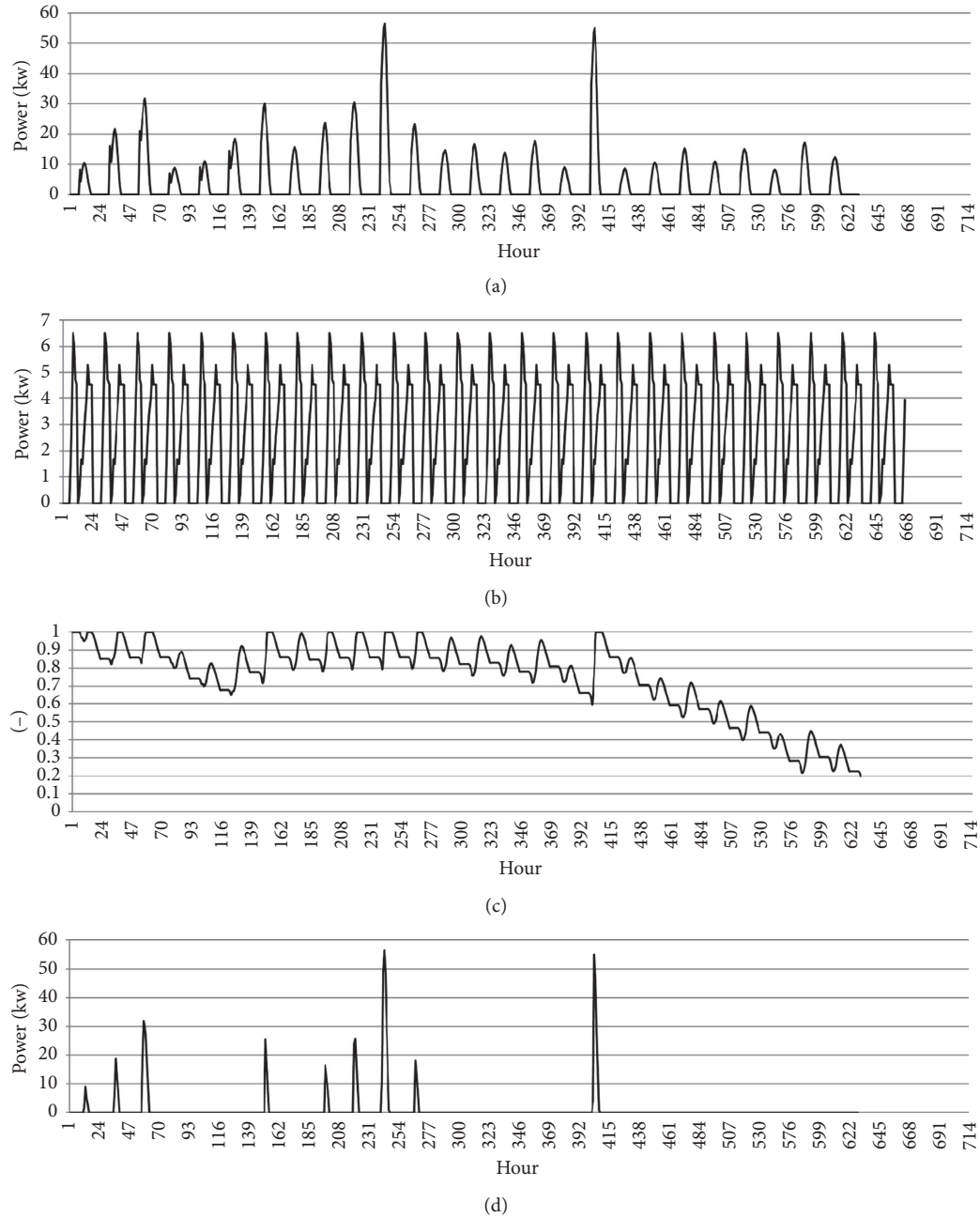
FIGURE 8: Designed system hourly performance for a failed case for June, with $d = 3$, in terms of the delivered energy E_{PV}, load demand $Q_{L,h}$, excess energy curtailed E_{loss}, and state of charge of the batteries SOC. (a) E_{PV}, (b) Q_L, (c) SOC, and (d) E_{LOSS}.

are shown in Table 7, including both critical and noncritical operation. Although the values are not identical, these success rates assure a minimum of 95% when operating with noncritical loads.

8. Discussion

The results obtained by adopting the proposed methodology reveal substantive differences with respect to the forecast made using the conventional approach. However, it is necessary to discuss the assumptions and considerations

made throughout the research to elucidate the causes of this difference.

Regarding the estimation of solar radiation, the same period was considered both for the adjustment using probability distributions and for determining the Kolmogorov–Smirnov statistic to measure the goodness of the fit. On the other hand, the number of parameters of the different distributions was not considered when performing the statistical test. This can result in overfitting based on the period analyzed, affecting its adaptability to estimate or forecast more recent data.

TABLE 7: Success rates SR% for all months of the year in order to achieve at least 95% and 99% of system autonomy.

	d	$P_{m,cor}$ (kWp)	$C_{L,cor}$ (kAh)	SR%											
				Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Noncritical loads	3	97.07	2.31	100	100	100	100	96	97	99	100	100	100	100	99
Critical loads	4	106.17	2.52	100	100	100	100	99	100	100	100	100	100	100	100

The method considered for estimating radiation is classified into long-term monthly average of daily solar radiation. According to the research carried out by [20], this type of estimate can result in lower estimation errors for the contemplated period, although the superiority and usability of the model are questionable. It is mainly because of the estimation error that a relatively stable and long enough period was considered in the radiation data to avoid the problems that the method itself entails. The aforementioned one may partially explain the difference between the results obtained versus conventional methodology. Another additional aspect to consider is the uncertainty that the use of meteorological data from sources such as NASA or METEONORM can provide, according to [40], precisely because the radiation data have not been routinely observed in most of the world's meteorological stations due to the high cost of instruments and technical requirements.

In the design phase of the dairy, it is necessary to mention that the proposed scenario combines the month with the lowest average daily global radiation according to the data provided, with the summer month with peak consumption. Therefore, the capacity required to meet the energy demand is being overestimated. The scenario, therefore, delivers a solution that is overestimating the storage capacity and thus becoming suboptimal, at least for the critical month, to meet the reliability standards at the lowest cost.

As in the research proposed in [6], this case considers a daily estimation method for solar radiation, but the main difference lies in the update of the different state variables involved to measure the compliance of reliability standards. In this sense, while the base method has a daily scale, the method proposed in this research updates the state variables and decomposes the global solar radiation considering an hourly scale. Furthermore, the extended method is not validated by a real implementation of the case study, not being able to determine which sizing method is more accurate.

However, given the characteristics of the sizing method, its result has to be contrasted with sophisticated and highly accurate models for the estimation and designing phase in order to overcome the limitations previously analyzed. In this sense, another aspect that can be expanded is the consideration of stochastic components for solar radiation on an hourly scale. The conditions for the design of this case study may also consider the possibility of simulating the energy supplied by the PV system with the estimated load profile for each month of the year, thus avoiding capacity overestimation.

One of the characteristics of the sizing of autonomous photovoltaic systems is the compliance with certain reliability levels; in this sense, it is possible to consider for future research the incorporation of a multiobjective function that considers not only the storage capacity, but also the implementation costs and other relevant performance indicators, as well as the incorporation of budget or size constraints for the installation of PV modules.

Given the current world energy scenario, it is now desirable to think of different sustainable alternatives for electricity generation in such a way as to minimize the environmental impact. Solar energy is widely available, and its use is increasing enormously, which leads to new challenges in improving the efficiency of PV systems and balancing supply and demand. If some loads can be shifted from night to day, better results can be obtained. Thus, the concept of load management becomes relevant to sizing standalone PV systems. In this sense, dynamic load management functionality can be introduced. In more recent studies, the effect of load profile uncertainty on the off-grid PV systems optimum design is analyzed, highlighting the importance when defining load profiles for off-grid PV-battery systems.

Finally, it is therefore recommended to perform an hourly, or even a more detailed simulation (e.g., simulation at minutely resolution), to have better control of the load profile and the power output fluctuations, clearly identifying the time when peak demand occurs, and utilize this valuable information for getting optimal results.

9. Conclusions

The results obtained from the stochastic simulation model show that the proposed standalone PV system can successfully meet the dairy's load demand at specified reliability, achieving a reduction of 44% in installed peak power and battery storage capacity for noncritical operation, and an 85% reduction for critical operation, as compared to conventional methodology, which is consistent with the results obtained in base research, thus validating the hourly-based method for a different geographical location.

Appendix

A. Nomenclature Used, Based on Kaplanis and Kaplani [6]

Nomenclature

AUX: auxiliary variable
a: lower bound of beta distribution
b: upper bound of beta distribution

$C_{\text{bat-ag}}$: for the ageing of the battery system due to cycling

$C_{\text{bat-c}}$: the efficiency of the batteries during charging process

$C_{\text{bat-d}}$: efficiency of the batteries during the discharging process

C_{charger}: efficiency of the charger

C_{inverter}: efficiency of the inverter

C_{L}: battery storage capacity (kAh)

$C_{\text{L,cor}}$: C_{L} corrected (kAh)

$C_{\text{pv-ag}}$: for the effect of the ageing of the PV

C_{Tc}: correction term for the temperature effect on PV efficiency

DE: the remaining amount of energy after satisfying all load requirements

DOD: depth of discharge of battery (dimensionless)

d: number of days of autonomy of a PV system

d_{cr}: number of days of autonomy of a PV system operating with critical loads

$d_{\text{n-cr}}$: number of days of autonomy of a PV system operating with noncritical loads

E_{loss}: energy dissipated due to excess energy and fully charged battery (kWh)

E_{PV}: energy delivered by the PV system (kWh)

F: loss of energy between the route of the PV generator and load demand (dimensionless)

F': loss of energy between the route of the battery bank and loads (dimensionless)

G: irradiance under standard test conditions (STC) (1 kW/m^2)

H: daily global solar radiation on the horizontal (kWh/m^2)

H_{m}: mean value of H (kWh/m^2)

I: hourly global solar radiation on the horizontal in a day n_j (kWh/m^2)

I_{b}: hourly direct beam radiation

I_{d}: hourly diffuse radiation

I_{ext}: hourly extraterrestrial radiation on a horizontal surface for an hour period

k: shape parameter of Gen. Extreme Value distribution

k_{T}: hourly clearness index

NOCT: nominal operating cell temperature (°C)

P_{m}: peak power (kWp)

$P_{\text{m,cor}}$: P_{m} corrected (kWp)

PSH: Peak Sun Hours (h/day)

PSH$_{\text{m}}$: mean value of PSH (h/day)

PSH$_{\text{min}}$: minimum value of PSH (h/day)

Q_{L}: total loads in a day (Wh/day)

$Q_{\text{L,h}}$: load demand at any time (kWh)

r_h: ratio of hourly global solar radiation to daily global solar radiation (dimensionless)

R_{m}: factor for converting the global solar radiation from the horizontal to the inclined plane of the PV panels, mean value for a specific month (dimensionless)

R: factor for converting the hourly solar radiation from the horizontal surface to the inclined plane, for a specific hour (dimensionless)

R_{b}: ratio of beam radiation on the tilted surface to that on a horizontal surface at any time

S_{o}: length of the day (h)

SOC: state of charge of battery % (dimensionless)

SOC$_{\text{a}}$: the remaining SOC carried onto the next hour (dimensionless)

TC: temperature coefficient

T_{a}: maximum average ambient temperature in the daytime for the sizing month (°C)

T_{C}: cell temperature (°C)

V: DC transfer voltage. Also, nominal battery operating voltage (V)

Greek letters

α: scale parameter of Wakeby distribution; shape parameter of Weibull distribution

α_1, α_2: shape parameters of beta distribution

β: shape parameter of Wakeby distribution; scale parameter of Weibull distribution; slope

γ: shape parameter of Johnson SB distribution; scale parameter of Wakeby distribution

δ: declination; shape parameter of Johnson SB distribution; shape parameter of Wakeby distribution

λ: scale parameter of Johnson SB distribution

ξ: location parameter of Johnson SB and Wakeby distributions

μ: location parameter of Gen. Extreme Value and Normal distributions

ρ_{g}: albedo

σ: scale parameter of Normal and Gen. Extreme Value distributions

σ_{H}: standard deviation of H

σ_{G}: standard deviation of the Gaussian curve

ϕ: site latitude

ω: hour angle

B. Flowchart of the Stochastic Proposed Methodology, Based on Kaplani and Kaplanis [6]
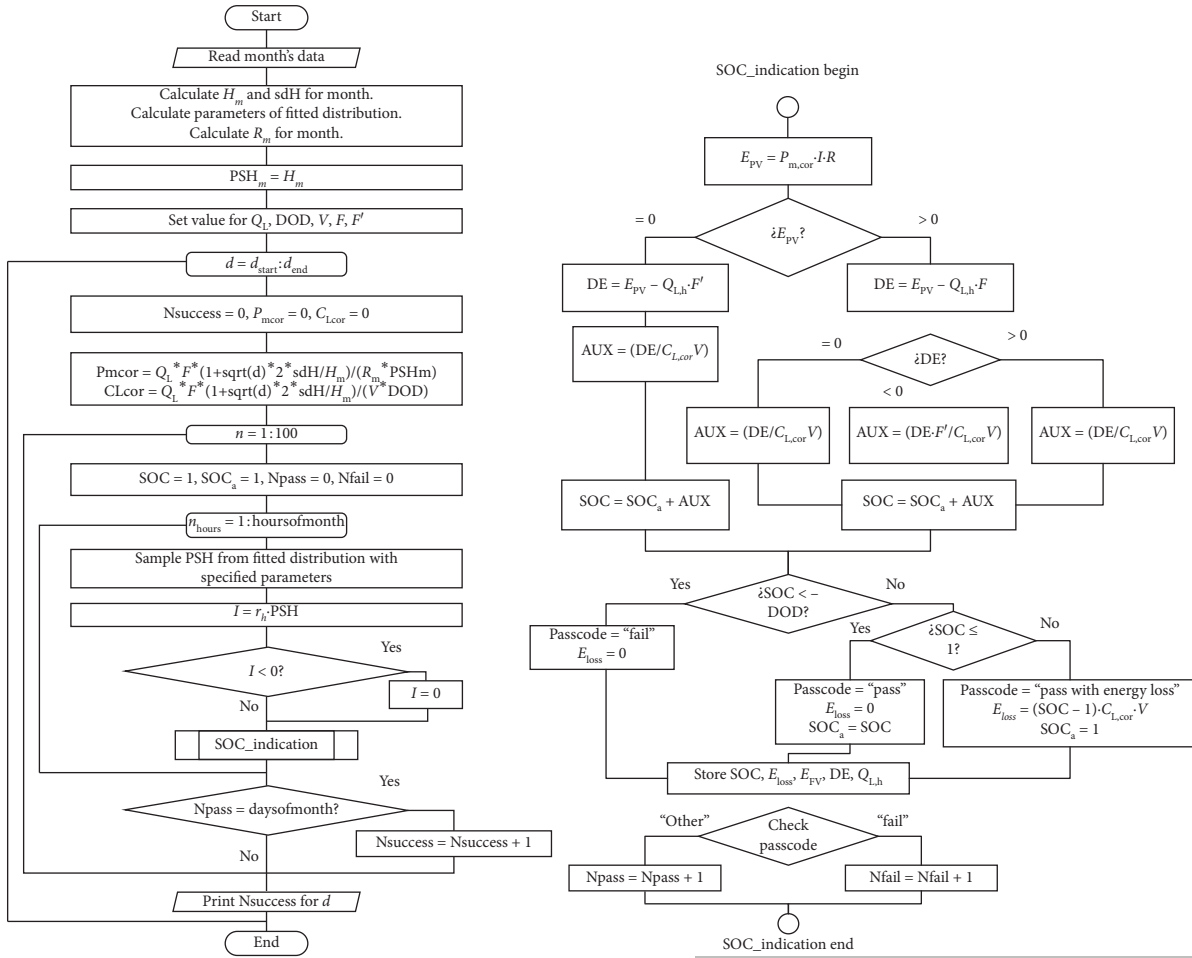
C. Maximum Log-Likelihood Estimation

To adjust the distributions for the subsequent comparison of the goodness of fit, the parametric estimation of the distributions is performed using the maximum likelihood estimation (MLE). Considering the set $\{X_i, i = 1, \ldots, n\}$ of n independent and identically distributed random variables, taken from a continuous probability distribution characterized by the distribution parameter θ, and representing its probability density function as $f(x|\theta)$, the likelihood function L is given by the following joint probability function P:

$$L(x_1, x_2, \ldots, x_n|\theta) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n|\theta), \tag{C.1}$$

where the realization of said random variables, that is, the observations, is represented by the set $\{x_i\}_n$. Developing the previous expression and considering the condition of random variables (iid), the likelihood function is given by

$$L(x_1, x_2, \ldots, x_n|\theta) = f(x_1, x_2, \ldots, x_n|\theta) = f(x_1|\theta)f(x_2|\theta), \ldots, f(x_n|\theta). \tag{C.2}$$

Therefore, the maximum likelihood estimator $\widehat{\theta}_{ML}$ is the estimator of the parameter θ that maximizes the value of this joint probability function. That is,

$$L(x_1, x_2, \ldots, x_n|\widehat{\theta}_{ML}) = \max_{\theta} L(x_1, x_2, \ldots, x_n|\theta). \tag{C.3}$$

The above sentence applies without loss of generality for the parameter vector $\theta = (\theta_1, \theta_2, \ldots, \theta_m)$.

D. Kolmogorov–Smirnov Test

The Kolmogorov–Smirnov (K-S) test is commonly used to decide if a sample from a population comes from a determined continuous probability distribution, from which its specification parameters are known. The test is defined by the hypotheses:

H_0: the sample follows a specified distribution

H_a: the sample does not follow a specified distribution

And the K-S statistic D_n is defined as

$$D_n = \sup_x |F_n(x) - F(x)|, \tag{D.1}$$

where $F_n(x)$ is the empirical distribution and $F(x)$ is the theoretical cumulative distribution function calculated over ordered observations $\{X_i, i = 1, \ldots, n\}$ of the sample. This value is compared with the K-S tables of two-tailed critical values $D_{\mathrm{cr},\alpha}$ for a given significance percentage α, which usually takes a value of 5%. Therefore, the null hypothesis is rejected when the condition $D_n > D_{\mathrm{cr},\alpha}$ is satisfied. To quantify this decision, avoiding the dependence of the significance value set a priori by the researcher to reject the null hypothesis, the p value is calculated to choose between the fitted distributions.

Data Availability

The information used in this study is referenced, respectively, and the data from the case study are obtained from a dairy in southern Chile.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. W. Sun, "Changes in energy consumption and energy intensity: a complete decomposition model," *Energy Economics*, vol. 20, no. 1, pp. 85–100, 1998.

[2] G. P. Peters, G. Marland, C. Le Quéré, T. Boden, J. G. Canadell, and M. R. Raupach, "Rapid growth in CO_2 emissions after the 2008-2009 global financial crisis," *Nature Climate Change*, vol. 2, no. 1, pp. 2–4, 2012.

[3] D. Watts, M. F. Valdés, D. Jara, and A. Watson, "Potential residential PV development in Chile: the effect of net metering and net billing schemes for grid-connected PV systems," *Renewable and Sustainable Energy Reviews*, vol. 41, pp. 1037–1051, 2015.

[4] A. Chouder and S. Silvestre, "Automatic supervision and fault detection of PV systems based on power losses analysis," *Energy Conversion and Management*, vol. 51, no. 10, pp. 1929–1937, 2010.

[5] M. Roger and V. Jerry, *Photovoltaic Systems Engineering*, CRC Press, Boca Raton, FL, USA, 2000.

[6] E. Kaplani and S. Kaplanis, "A stochastic simulation model for reliable PV system sizing providing for solar radiation fluctuations," *Applied Energy*, vol. 97, pp. 970–981, 2012.

[7] M. Breen, J. Upton, and M. D. Murphy, "Photovoltaic systems on dairy farms: financial and renewable multi-objective optimization (FARMOO) analysis," *Applied Energy*, vol. 278, p. 115534, 2020.

[8] M. De Blas, J. Appelbaum, J. L. Torres, A. García, E. Prieto, and R. Illanes, "A refrigeration facility for milk cooling powered by photovoltaic solar energy," *Progress in Photovoltaics: Research and Applications*, vol. 11, no. 7, pp. 467–479, 2003.

[9] J. Page, "The estimation of monthly mean values of daily total shortwave radiation on vertical and inclined surfaces from sunshine records for latitude 40°N-40°S," in *Proceedings of the UN Conference on New Sources of Energy*, pp. 378–390, Rome, Italy, August 1961.

[10] A. Baig, P. Akhter, and A. Mufti, "A novel approach to estimate the clear day global radiation," *Renewable Energy*, vol. 1, no. 1, pp. 119–123, 1991.

[11] T. Khatib, I. A. Ibrahim, and A. Mohamed, "A review on sizing methodologies of photovoltaic array and storage battery in a standalone photovoltaic system," *Energy Conversion and Management*, vol. 120, pp. 430–448, 2016.

[12] S. Alsadi and T. Khatib, "Photovoltaic power systems optimization research status: a review of criteria, constrains, models, techniques, and software tools," *Applied Sciences*, vol. 8, no. 10, p. 1761, 2018.

[13] M. M. Fouad, L. A. Shihata, and E. I. Morgan, "An integrated review of factors influencing the performance of photovoltaic panels," *Renewable and Sustainable Energy Reviews*, vol. 80, pp. 1499–1511, 2017.

[14] C. J. Lanigan, T. Donncilan, K. Hanrahan et al., *An Analysis of Abatement Potential of Greenhouse Gas Emissions in Irish Agriculture 2021–2030*, Teagasc, Carlow, Ireland, 2018.

[15] M. Breen, J. Upton, and M. D. Murphy, "Photovoltaic systems on dairy farms: an economic analysis," in *Proceedings of the Annual International Meeting*, Detroit, MI, USA, July 2018.

[16] Food and Agriculture Organization, *Outlook 2018–2027 Dairy and Dairy Products 2018*, Food and Agriculture Organization, Rome, Italy, 2018, http://www.agrioutlook.org/commodities/Agricultural-Outlook-2018-Dairy.pdf.

[17] J. Bruinsma and N. Alexandratos, *World Agriculture Towards 2030/2050: The 2012 Revision,* Food and Agriculture Organization, Rome, Italy, 2012, http://www.fao.org/docrep/016/ap106e/ap106e.pdf.

[18] P. Shine, J. Upton, P. Sefeedpari, and M. D. Murphy, "Energy consumption on dairy farms: a review of monitoring, prediction modelling, and analyses," *Energies*, vol. 13, no. 5, p. 1288, 2020.

[19] C. Zhang, P. Campana, J. Yang, J. Zhang, and J. Yan, "Can solar energy be an alternative choice of milk production in dairy farms?—A case study of integrated PVWP system with alfalfa and milk production in dairy farms in China," *Energy Procedia*, vol. 105, pp. 3953–3959, 2017.

[20] J. Zhang, L. Zhao, S. Deng, W. Xu, and Y. Zhang, "A critical review of the models used to estimate solar radiation," *Renewable and Sustainable Energy Reviews*, vol. 70, pp. 314–329, 2017.

[21] A. Angstrom, "Solar and terrestrial radiation. Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation," *Quarterly Journal of the Royal Meteorological Society*, vol. 50, no. 210, pp. 121–126, 1924.

[22] K. Bakirci, "Models of solar radiation with hours of bright sunshine: a review," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 9, pp. 2580–2588, 2009.

[23] F. Besharat, A. A. Dehghan, and A. R. Faghih, "Empirical models for estimating global solar radiation: a review and case study," *Renewable and Sustainable Energy Reviews*, vol. 21, pp. 798–821, 2013.

[24] J. Qin, Z. Chen, K. Yang, S. Liang, and W. Tang, "Estimation of monthly-mean daily global solar radiation based on MODIS and TRMM products," *Applied Energy*, vol. 88, no. 7, pp. 2480–2489, 2011.

[25] A. S. Al-Sumaiti, M. H. Ahmed, S. Rivera, M. S. El Moursi, M. M. A. Salama, and T. Alsumaiti, "Stochastic PV model for power system planning applications," *IET Renewable Power Generation*, vol. 13, no. 16, pp. 3168–3179, 2019.

[26] C. Jung and D. Schindler, "Global comparison of the goodness-of-fit of wind speed distributions," *Energy Conversion and Management*, vol. 133, pp. 216–234, 2017.

[27] Nasa, Surface Meteorology and Solar Energy, Sol Energy 2011, http://org/commodities/Agricultural-Outlook-2018-Dairy.pdf.

[28] S. A. Klein, "Calculation of monthly average insolation on tilted surfaces," *Solar Energy*, vol. 19, no. 4, pp. 325–329, 1977.

[29] J. A. Duffie and W. A. Beckman, *Solar Engineering of Thermal Processes*, Wiley, Hoboken, NJ, USA, 4th edition, 2013.

[30] D. S. Dimitrova, V. K. Kaishev, and S. Tan, "Computing the Kolmogorov-smirnov distribution when the underlying CDF is purely discrete, mixed, or continuous," *Journal of Statistical Software*, vol. 95, no. 10, pp. 1–42, 2020.

[31] G. Marsaglia, W. W. Tsang, and J. Wang, "Evaluating Kolmogorov's distribution," *Journal of Statistical Software*, vol. 8, no. 18, pp. 1–4, 2003.

[32] P. C. Jain, S. Jain, and C. F. Ratto, "A new model for obtaining horizontal instantaneous global and diffuse radiation from the daily values," *Solar Energy*, vol. 41, no. 5, pp. 397–404, 1988.

[33] S. N. Kaplanis, "New methodologies to estimate the hourly global solar radiation; comparisons with existing models," *Renewable Energy*, vol. 31, no. 6, pp. 781–790, 2006.

[34] B. Y. H. Liu, R. C. Jordan, S. Dongaonkar, and M. A. Alam, "The long-term average performance of flat-plate solar-energy collectors," *Solar Energy*, vol. 7, no. 2, pp. 53–74, 1963.

[35] H. C. Hottel and B. B. Woertz, "The performance of flat-plate solar–heat collectors," *Transactions of the American Society of Mechanical Engineers*, vol. 64, pp. 91–104, 1942.

[36] D. G. Erbs, S. A. Klein, and J. A. Duffie, "Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation," *Solar Energy*, vol. 28, no. 4, pp. 293–302, 1982.

[37] A. Luque and S. Hegedus, *Handbook of Photovoltaic Science and Engineering*, Wiley, Hoboken, NJ, USA, 2nd edition, 2011.

[38] N. Lukač, S. Seme, D. Žlaus, G. Štumberger, and B. Žalik, "Buildings roofs photovoltaic potential assessment based on LiDAR (light detection and ranging) data," *Energy*, vol. 66, pp. 598–609, 2014.

[39] A. Green, "Life cycle costing for batteries in telecom applications," in *Proceedings of the INTELEC—Twentieth International Telecommunications Energy Conference (Cat. No.98CH36263)*, pp. 1–7, San Francisco, CA, USA, October 1998.

[40] L. Wang, O. Kisi, M. Zounemat-Kermani, G. A. Salazar, Z. Zhu, and W. Gong, "Solar radiation prediction using different techniques: model evaluation and comparison," *Renewable and Sustainable Energy Reviews*, vol. 61, pp. 384–397, 2016.