



Revista Facultad de Ingeniería

ISSN: 0121-1129

ISSN: 2357-5328

Universidad Pedagógica y Tecnológica de Colombia

Pachajoa, Dalila-Mercedes; Mora-Paz, Héctor-Andrés; Mayorca-Torres, Dagoberto  
Comparison of Kernel Functions in the Classification  
of Irradiance Zones from Multispectral Satellite Images

Revista Facultad de Ingeniería, vol. 30, no. 58, e106, 2021, October-December  
Universidad Pedagógica y Tecnológica de Colombia

DOI: <https://doi.org/10.19053/01211129.v30.n58.2021.13845>

Available in: <https://www.redalyc.org/articulo.oa?id=413969810007>

- How to cite
- Complete issue
- More information about this article
- Journal's webpage in redalyc.org

redalyc.org

Scientific Information System Redalyc

Network of Scientific Journals from Latin America and the Caribbean, Spain and  
Portugal

Project academic non-profit, developed under the open access initiative

**Revista Facultad de Ingeniería**

Journal Homepage: <https://revistas.uptc.edu.co/index.php/ingenieria>



# Comparison of Kernel Functions in the Classification of Irradiance Zones from Multispectral Satellite Images

Dalila-Mercedes Pachajoa<sup>1</sup>

Héctor-Andrés Mora-Paz<sup>2</sup>

Dagoberto Mayorca-Torres<sup>3</sup>

**Received:** October 11, 2021

**Accepted:** December 16, 2021

**Published:** December 20, 2021

**Citation:** D.-M. Pachajoa, H.-A. Mora-Paz, D. Mayorca-Torres, "Comparison of Kernel Functions in the Classification of Irradiance Zones from Multispectral Satellite Images," *Revista Facultad de Ingeniería*, vol. 30 (58), e13845, 2021. <https://doi.org/10.19053/01211129.v30.n58.2021.13845>

## Abstract

Due to the growing energy demand and the eminent global warming, there is special interest in the prediction of irradiance based on the reflectance obtained from satellites such as NASA Landsat, since it allows to know where it is more efficient to

<sup>1</sup> Universidad de Nariño (Pasto-Nariño, Colombia). [dalila@udenar.edu.co](mailto:dalila@udenar.edu.co). ORCID: <https://orcid.org/0000-0002-4096-7609>

<sup>2</sup> M. Sc. Universidad Cesmag (Pasto-Nariño, Colombia). [hamora@unicesmag.edu.co](mailto:hamora@unicesmag.edu.co). ORCID: <https://orcid.org/0000-0003-3097-4757>

<sup>3</sup> Universidad Mariana (Pasto-Nariño, Colombia). ORCID: <https://orcid.org/0000-0002-4342-0238>



place photovoltaic receivers. Although there are studies for obtaining regression models with alternative Kernel functions, their performance for classification models is unknown and it is here where this research focuses. The study couples alternative Kernel functions to the support vector machines (SVM) algorithm for classification problems, where the best configuration for these algorithms is explored to finally obtain a set of irradiance maps zoned by class.

**Keywords:** classification; Kernel functions; Landsat; multispectral satellite images; photovoltaic energy; Support Vector Machines.

### **Comparativo de funciones Kernel en la clasificación de zonas de irradiancia a partir de imágenes satelitales multiespectrales**

#### **Resumen**

Debido a la creciente demanda de energía y al eminente calentamiento global, existe especial interés en la predicción de irradiancia basada en la reflectancia obtenida de satélites como el Landsat de la NASA, ya que permite saber dónde es más eficiente colocar receptores fotovoltaicos. Si bien existen estudios para la obtención de modelos de regresión con funciones Kernel alternativas, se desconoce su desempeño para modelos de clasificación, y es aquí donde se enfoca esta investigación. El estudio combina funciones de Kernel alternativas al algoritmo máquinas de soporte vectorial (SVM) para problemas de clasificación, donde se explora la mejor configuración para estos algoritmos, y así finalmente obtener un conjunto de mapas de irradiancia zonificados por clase.

**Palabras claves:** clasificación; energía fotovoltaica; funciones Kernel; imágenes satelitales multiespectrales; Landsat; máquinas de soporte vectorial.

### **Comparação de funções de kernel na classificação de zonas de irradiância de imagens de satélite multiespectrais**

#### **Resumo**

Devido à crescente demanda de energia e ao eminente aquecimento global, há especial interesse na previsão da irradiância a partir da refletância obtida de satélites como o NASA Landsat, pois permite saber onde é mais eficiente colocar

os receptores fotovoltaicos. Embora existam estudos para obtenção de modelos de regressão com funções alternativas do Kernel, seu desempenho para modelos de classificação é desconhecido e é aqui que se concentra esta pesquisa. O estudo acopla funções Kernel alternativas ao algoritmo de máquinas de vetores de suporte (SVM) para problemas de classificação, onde a melhor configuração para esses algoritmos é explorada para finalmente obter um conjunto de mapas de irradiância zoneados por classe.

**Palavras-chave:** classificação; energia fotovoltaica; imagens de satélite multiespectrais; Kernel functions; Landsat; máquinas de vetor de suporte.

## I. INTRODUCTION

The abstraction of reality has long become an essential means of thinking and expressing, being indispensable in reasoning, knowledge, and interrelation. Abstraction has been spread in terms of machines (Machine Learning), where reality is abstracted through a careful analysis of all variables, to the point that the machine is able to predict based on the incoming data. Support Vector Machines (SVM) is one of the most widely used algorithms to perform prediction tasks in different scenarios. One of the main features that make it a good predictor is the possibility of finding nonlinear patterns in the data through the Kernel trick, which takes the original feature space where the data are not linearly separable to an infinite dimensional Hilbert space where the data are linearly separable [1]. In that sense, traditional Kernel functions such as linear, polynomial, and Gaussian (rbf) Kernel have been used. However, it is possible to improve the results of the algorithms using alternative kernel functions [2].

In 2040, total energy demand in the world will increase by 30% and most of this consumption will come from developing countries. In addition, 37% of electricity generation is expected to come from renewable sources, especially wind and solar. So then, solar energy is retaken, allowing electricity to be obtained completely free and renewable, in this sense, photovoltaic solar energy is the electrical energy generated by the photovoltaic effect that occurs when solar radiation falls on a photovoltaic panel[3]. Hence, there is a clear worldwide need to maximize the use of photovoltaic solar energy, which main advantage is that no greenhouse gases or pollutants are emitted. In addition, this energy can be used anywhere in the world, reaching remote places and isolated homes where the power lines do not extend to, without leaving behind that solar energy will irradiate the earth for millions of years; in fact, it is considered one of the most efficient renewable technologies in the fight against climate change. In the last decade, photovoltaic solar energy has experienced a drastic reduction in costs that has made it, along with wind power, one of the most promising energy technologies for the future. Thus, since the installed photovoltaic capacity in the world stood at 495 GW at the end of 2018, the

International Energy Agency forecasts that by 2040 it will have increased sixfold, to over 3,000 GW (and even 4,800 MW in its most sustainable scenario) [4].

In relation to this, in Colombia, entities such as the Mining and Energy Planning Unit (Unidad de Planeación Minero Energética - UPME) and the Institute for Planning and Promotion of Energy Solutions for Non-Interconnected Zones (Instituto de Planificación y Promoción de Soluciones Energéticas para Zonas No Interconectadas – IPSE) have identified development initiatives in rural regions of the country, with projects such as the sustainable rural energization plan (Plan de Energización Rural Sostenible - PERS), allowing with the analysis of energy information, the construction of documents on the topics such as energy supply and demand, and energy policy guidelines [4], [5]. Specifically, in the department of Nariño, the oil deficit produces fuel shortages for transportation and areas without continuous electricity supply [6]. From PERS also comes the project Analysis of Energy Opportunities (ALTERNAR), in which —with an extrapolation model— it was found that ANN and SVM (for regression) achieved the best results in the prediction of irradiance from Landsat and MODIS satellite images. Based on this study, Mora [7] improves the previous results using alternative Kernel functions. Under this background, adjustments to experiment with an SVM for classification with alternative kernel functions were made, obtaining a classification model from the discretization of irradiance, the best Kernel function is rbf, and the radial basic function. Finally, comparisons of each Kernel function and geographic visualizations were obtained from the inferences given by the best pipeline; thus, having less variability in the classification model, irradiance zones can be identified more easily. In conclusion for to generate irradiance classification models since Landsat dataset it is recommended use a discretization by equal ranks where applying SVM algorithm with radial basic or rational quadratic kernel, with max-min normalizer.

## II. METHODOLOGY

The following sections show the tools used for the development of this research, and how the experiments were set up to obtain each product.

## A. Materials and Methods

The data of this study correspond to multispectral data taken from the Nasa Landsat sensor for the department of Nariño, Colombia, which have been preprocessed by [8] and [7]. It consists of 434 records, the feature vector corresponds to geolocation data (latitude and longitude) and 7 spectral bands, and the target variable (value) corresponds to the irradiance. Table 1 describes each of the variables and Figure 1 shows the phases of the process.



Fig. 1. Phases of the process.

Table 1. Description of the data set.

Variable	Description	Range	Units
Latitude	Latitude	-8789850.0, -8554950.0	Meters
Longitude	Length	45000.0, 294750.0	Meters
Band1	Coastal Band/Aerosol	0.43,0.45	Micro meters
Band2	Blue Band	0.45,0.51	Micro meters
Band3	Green band	0.53,0.59	Micro meters
Band4	Red Band	0.64,0.67	Micro meters
Band5	Near Infrared Band NIR	0.85,0.88	Micro meters
Band6	Shortwave infrared band SWIR	1.57,1.85	Micro meters
Band7	Thermal infrared band	2.11,2.99	Micro meters
Value	Irradiance	188.5, 247.3	$W/m^2$

The geographic coordinates of this dataset are projected in Mercator 3857 as well as the polygon named narino\_3857.shp, which was used for geospatial visualization of the predictions. The experiments developed in this study were written using Python language on Google Colab.

## B. Coupling Kernel Functions to Sklearn

To obtain the best configuration for the SVM algorithm, we initially coupled new kernel functions to it by extending the SVC implementation of the sklearn library for classification as described in more detail in [7], providing the SVM with the rational

quadratic, truncated, canberra, radial basic, triangle, and hyperbolic functions. Table 2 shows the mathematical definition of the above kernel functions.

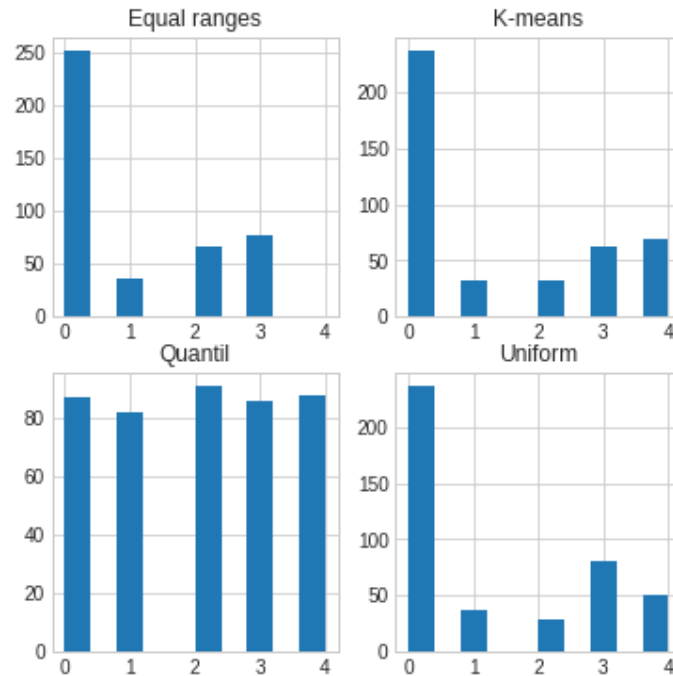
**Table 2.** Kernel functions coupled to sklearn. Source [2].

Kernel functions	Formal mathematical definition
<i>Rational quadratic (rq)</i>	$k^{rq}(\mathbf{x}, \mathbf{x}') = 1 - \frac{\ \mathbf{x} - \mathbf{x}'\ ^2}{\ \mathbf{x} - \mathbf{x}'\ ^2 + a}, a > 0$
<i>Truncated (tru)</i>	$k^{tru}(\mathbf{x}, \mathbf{x}') = \frac{1}{d} \sum_{i=1}^d \max \left\{ 0, 1 - \frac{ x_i - x'_i }{\gamma_i} \right\}, \gamma_i > 0$
<i>Canberra (can)</i>	$k^{can}(\mathbf{x}, \mathbf{x}') = 1 - \frac{1}{d} \sum_{i=1}^d \gamma_i \frac{ x_i - x'_i }{x_i + x'_i}, \gamma_i \in (0, 1]$
<i>Radial basic (rb)</i>	$k^{rb}(\mathbf{x}, \mathbf{x}') = \left( \sum_{i=1}^d \exp \{ -\gamma_i (x_i - x'_i)^2 \} \right)^m, \gamma_i > 0, m \in \mathbb{N}$
<i>Triangle (tri)</i>	$k^{tri}(\mathbf{x}, \mathbf{x}') = \begin{cases} 1 - \frac{\ \mathbf{x} - \mathbf{x}'\ }{a} & \text{si } \ \mathbf{x} - \mathbf{x}'\  \leq a, a > 0 \\ 0 & \text{otherwise} \end{cases}$
<i>Hyperbolic (hyp)</i>	$k^{hyp}(\mathbf{x}, \mathbf{x}') = \tanh (a\langle \mathbf{x}, \mathbf{x}' \rangle + b), a > 0, b < 0$

### C. Data Exploration and Preprocessing

After coupling the kernel functions, the Landsat dataset was loaded and the irradiance (target variable) was converted to a discrete variable, for this purpose 4 different encoders were examined to discretize the variable, namely: equal ranges, k-means, quantiles, and uniform distribution. Figure 2 shows the discretization for 5 classes with the above-described techniques.





**Fig. 2.** Irradiance discretization with different techniques.

As shown in Figure 2., each discretizer offers a different data distribution, although it is advisable to have balanced classes to train a classifier such as the one offered by the discretization in quartiles in this research. We experimented with the 4 distributions, as explained in the next section.

#### ***D. Tuning of Hyperparameters***

For the comparison of each kernel function to be fair, we experimented using the 4 discretizers of the previous section, 3 different normalizers were used on the feature vector, these are: minimum maximum scaling (MinMax), uniform scaling or standardization (Std), and scaling to the vector norm (Norm). The data was then partitioned leaving 20% of the data for testing, and 80% for training using the 2021 seed (random state). Next, a random search engine was used, configured with a stratified validation with 5 folds. Each search was run on each kernel function (7 kernel functions: 6 from Table 3 plus the RBF function), for each data normalization (3 normalizers) and discretizer (4 discretizers), running a total of 112 searches. The values to be searched for the regularization coefficient (C) are in the logarithmic

space with lower bound 0, upper bound 5, consisting of 10 elements. The hyperparameter of each kernel function (coef0 for rational quadratic, gamma for the other kernel functions) is in the logarithmic space with lower bound -4, upper bound 4, consisting of 20 elements.

### ***E. Obtaining the Best Model***

Once the hyperparameter tuning was performed, the results were stored to evaluate the commitment of each configuration, in terms of accuracy and training and inference times. Subsequently, the best configurations were chosen for each data discretization and the accuracy of each model was evaluated with the test data. Finally, the extrapolated predictions of this model on all the data were extracted and the geographic visualization was performed, interpolating the data using the Kriging algorithm [9], and the visualization was compared with the maps generated in the state of the art for regression models on these same data.

The following section shows the results obtained.

## **III. RESULTS**

The products generated in this research are shown below. First, the coupling of the kernel functions for classification is shown, followed by the results obtained in the hyperparameter tuning stage, and finally, the maps generated with the best model found are shown.

### ***A. Coupling Kernel Functions for Classification***

To couple the kernel functions for classification, the SVC class of the Scikit-learn library was extended as shown in Figure 3. There, it is observed that to couple the alternative kernel functions it is necessary to extend the SVC class of Scikit-Learn, once this was done, the constructor of the class was overwritten and the Scikit-Learn Custom Kernel function procedure was used, which transforms the feature vector, using a gram matrix generated from the kernel function defined in the KernelF class. The KernelF class has been implemented by Mora [7].



**Fig. 3.** Coupling kernel functions in the SVC class of Scikit-Learn.

The above implementation can be installed in Python using the PIP ver command (<https://pypi.org/project/sklearnkernels/>) and can be contributed to by cloning the following repository (<https://github.com/magohector/sklearnkernels>).

### ***B. Search Results for Hyperparameter Tuning***

As mentioned in the hyperparameter tuning section, several searches were used to perform a fair comparison for each kernel function. For this purpose, 3 pipelines were structured to couple the normalizers and the KSV algorithm. The searches were then run and the results for each discretizer were stored in 3 different CSV files. Table 3 provides a better description of each of the resources obtained.

**Table 3.** Products obtained in the hyperparameter tuning.

Product	Description
KernelsClassifier.ipynb	Scripts for hyperparameter tuning and storage of the results obtained in the files equal.csv, kmeans.csv, quantil.csv and uniform.csv.
GetModel.ipynb	Scripts in charge of obtaining the comparative graphs, finding the best model with hold out, and generating the geographic visualization for the best configuration found.
equal.csv	Files containing the results of the searches for hyperparameter tuning, the name corresponds to the type of discretization applied, each file consists of the following columns: <ul style="list-style-type: none"> <li>Scaler: Type of standardization applied</li> <li>Kernel: Type of kernel function.</li> <li>mean_test_score: Average accuracy for 5 partitions.</li> <li>std_test_score: Standard deviation of accuracy for 5 partitions.</li> <li>mean_fit_time: Average training time for 5 partitions.</li> <li>std_fit_time: Standard deviation of training time for 5 partitions.</li> <li>mean_score_time: Average inference time for 5 partitions.</li> <li>std_score_time: Standard deviation of inference time for 5 partitions.</li> </ul>
kmeans.csv	
quantil.csv	
uniform.csv	

The files listed at Table 3 can be downloaded from the following repository <https://github.com/magohector/IrradianceClasification>. Table 4 shows the consolidated results of the files (.csv) listed, the results with accuracy greater than 0 have been filtered. Table 4 indicating the discretizer (Dis), the normalizer (Sca), the kernel function (kernel), the average accuracy (Mts), the standard deviation of accuracy (Sts), the average training time (Mft), the standard deviation of training time (Sft), the average inference time (Mst), and the standard deviation of inference time (Sst). Table 4 highlights the best results per discretizer and normalizer.

**Table 4.** Best results obtained in hyperparameters tuning.

Dis	Sca	kernel	Mts	Sts	Mft	Sft	Mst	Sst
Equal	SScaler	rq	0.7553	0.1182	2.1113	0.0310	0.4978	0.0064
		rbf	<b>0.7759</b>	<b>0.1398</b>	<b>0.0120</b>	<b>0.0026</b>	<b>0.0022</b>	<b>0.0012</b>
		tru	0.7255	0.0615	3.3716	0.0432	0.8520	0.0346
		can	0.7068	0.1237	3.9616	0.0767	0.9802	0.0317
		rb	0.7598	0.1321	2.8233	0.0440	0.6844	0.0223
		tri	0.7346	0.0944	1.8359	0.0348	0.4512	0.0105
		hyp	0.7598	0.1247	1.0527	0.0767	0.2848	0.0157
	MMScaler	rq	0.7598	0.1373	2.0953	0.0641	0.5191	0.0044
		rbf	0.7368	0.1093	<b>0.0079</b>	<b>0.0018</b>	<b>0.0017</b>	<b>0.0001</b>
		tru	0.7368	0.1134	3.4349	0.0305	0.8559	0.0209
		can	0.7231	0.0923	3.9946	0.0170	0.9586	0.0158
		rb	<b>0.7644</b>	<b>0.1289</b>	2.7902	0.0296	0.6889	0.0166
		tri	0.7414	0.1253	1.7964	0.0166	0.4396	0.0065
		hyp	0.7414	0.1109	0.9487	0.0147	0.2322	0.0024
	NMScaler	can	<b>0.7532</b>	<b>0.0569</b>	3.8465	0.0240	0.9494	0.0095
		tri	0.7000	0.1009	<b>1.6697</b>	<b>0.0231</b>	<b>0.4176</b>	<b>0.0106</b>
Kmean	Sscaler	rq	0.7208	0.1018	2.0316	0.0243	0.5029	0.0064
		rbf	0.7045	0.1242	<b>0.0132</b>	<b>0.0035</b>	<b>0.0024</b>	<b>0.0014</b>
		rb	<b>0.7254</b>	<b>0.1068</b>	2.7562	0.0143	0.6852	0.0117
		tri	0.7021	0.1496	1.7784	0.0099	0.4153	0.0592
	MMScaler	rq	<b>0.7187</b>	<b>0.0634</b>	2.0087	0.0109	0.4994	0.0068
		rbf	0.7047	0.0967	<b>0.0091</b>	<b>0.0017</b>	<b>0.0017</b>	<b>0.0001</b>
		tru	0.7021	0.1487	3.3643	0.0223	0.8269	0.0139
		tri	0.7021	0.1466	1.7994	0.0190	0.4496	0.0077
Uniform	Sscaler	can	<b>0.7072</b>	<b>0.0626</b>	<b>3.8854</b>	<b>0.0285</b>	<b>0.9606</b>	<b>0.0126</b>
		rq	0.7209	0.0662	2.0042	0.0165	0.4936	0.0060

Dis	Sca	kernel	Mts	Sts	Mft	Sft	Mst	Sst
		rbf	0.7253	0.1115	<b>0.0151</b>	<b>0.0045</b>	<b>0.0017</b>	<b>0.0001</b>
		tru	0.7026	0.0383	3.3016	0.0201	0.8276	0.0133
		rb	<b>0.7299</b>	<b>0.1109</b>	2.7998	0.0260	0.7034	0.0261
		tri	0.7138	0.0970	1.8126	0.0232	0.4254	0.0656
		hyp	0.7022	0.1144	0.9327	0.0143	0.2378	0.0054
	MMScaler	rq	<b>0.7279</b>	<b>0.0531</b>	2.0217	0.0119	0.5069	0.0241
		can	0.7072	0.0469	3.8420	0.0308	0.9589	0.0109
		rb	0.7051	0.0298	2.7294	0.0107	0.6849	0.0059
		tri	0.7184	0.1047	1.8248	0.0191	0.4544	0.0297
		hyp	0.7025	0.0513	<b>0.9317</b>	<b>0.0105</b>	<b>0.2315</b>	<b>0.0052</b>
	NMScaler	can	<b>0.7303</b>	<b>0.0301</b>	<b>3.8378</b>	<b>0.0217</b>	<b>0.9505</b>	<b>0.0102</b>

The accuracy of the searches performed in this hyperparameter configuration was also plotted and can be seen in Figure 4, which shows the accuracy (y axis), kernel function (x axis (namely: hyp, rb, rq, rbf, tru, tri, and can) and normalizer NMScaler (blue), Sscaler (Orange), NMScaler (Green) for each discretizer.

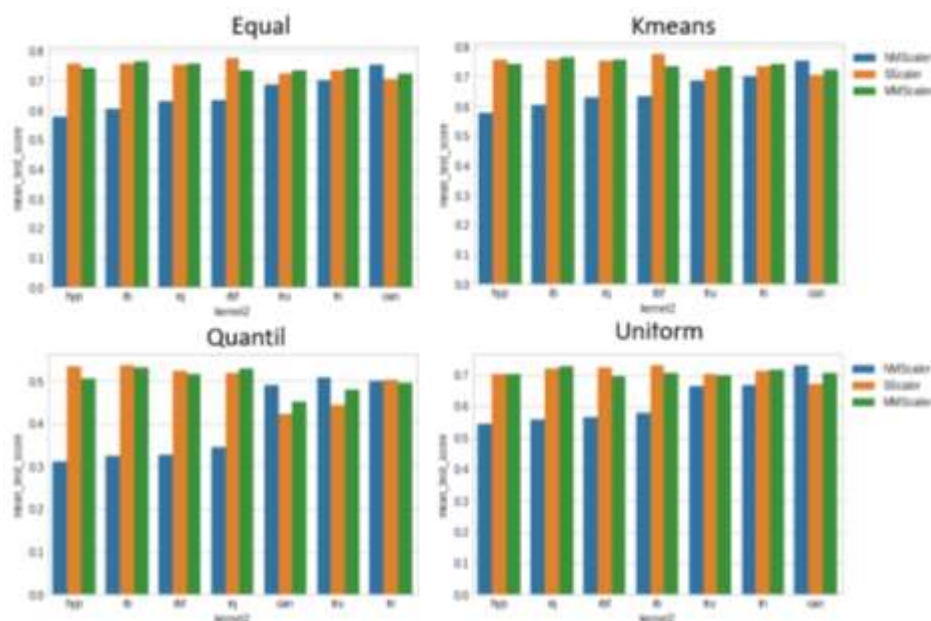
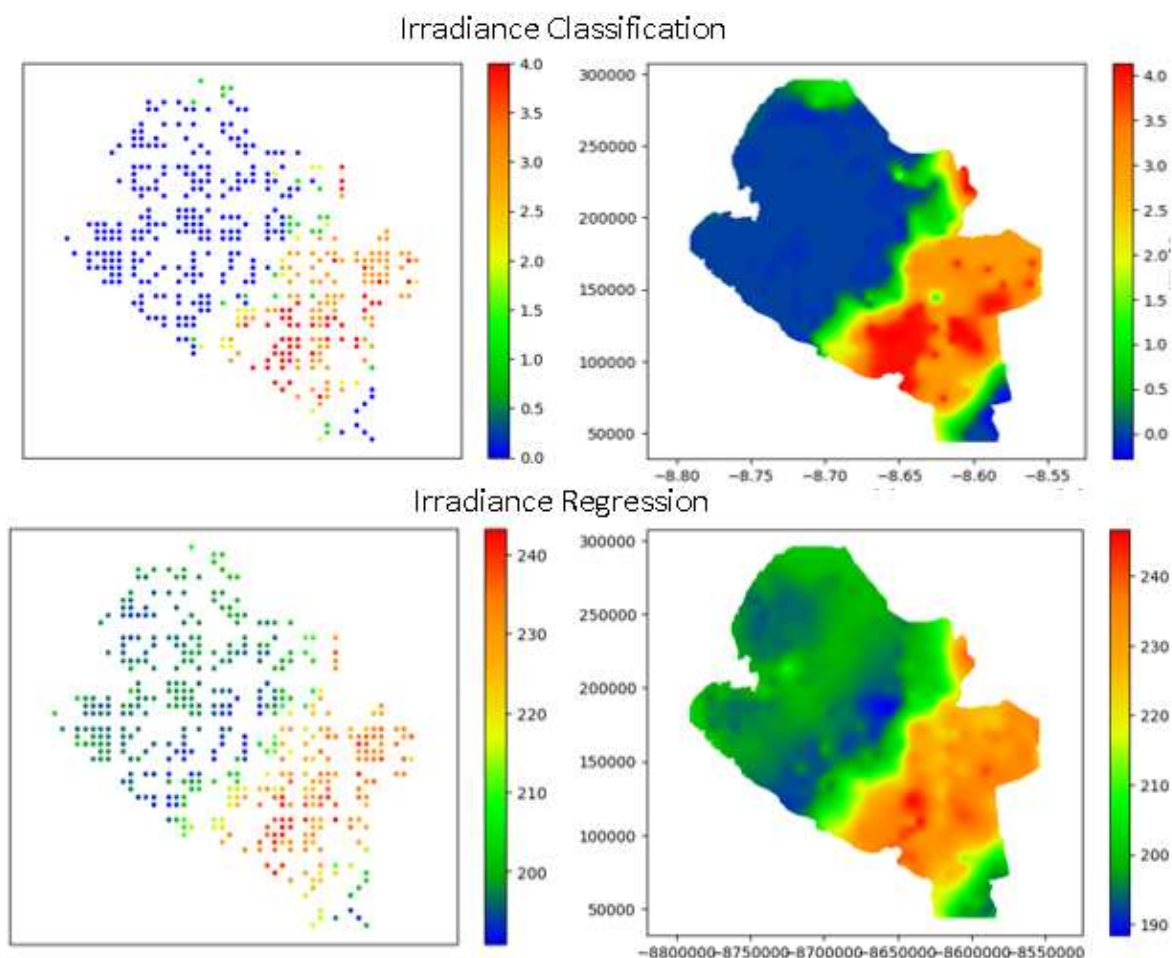


Fig. 4. Accuracy per discretizer in hyperparameter tuning.

The best models were taken from each discretizer and the results were evaluated with the test data, where the best configuration obtained an accuracy of 0.8161 for the model with equal discretizer, standard nomalizer, kernel function rbf, gamma

equal to 0.000695, and with regularization constant 2154.434690. With this model all the data were extrapolated and interpolated with the ordinary kriging algorithm with a spherical variogram, with a sample of 450 data. Figure 5 shows the data extrapolation, the interpolation of data with classification and the interpolation of state-of-the-art data.



**Fig. 5.** Extrapolation and interpolation with the classification and regression models.

The best models were taken from each discretizer and the results were evaluated with the test data, where the best configuration obtained an accuracy of 0.8161 for the model with equal discretizer, standard nomalizer, kernel function rbf, gamma equal to 0.000695, and with regularization constant 2154.434690. With this model all the data were extrapolated and interpolated with the ordinary kriging algorithm with an Spherical variogram, with a sample of 450 data. Figure 5 shows the data

extrapolation (mesh of points) and interpolation for the classification and regression problems, respectively. The same color scale has been used to generate the maps to facilitate visual comparison of the results obtained.

#### **IV. DISCUSSION**

As a starting point it is necessary to state that for the purposes of this study the accuracy gain was evaluated for 4 data discretizations, finding that the best way to discretize the irradiance to generate a classification model is the discretizer with equal ranks.

With respect to previous studies focused on the extrapolation of irradiance as a function of the ultraviolet, visible and infrared bands of the electromagnetic spectrum, a new alternative has been proposed, generating the extrapolation from a classification model; however, the accuracy of the model has only reached a value of approximately 0.82, compared to the determination coefficient of 0.94 [7]. Although they are different metrics, it is clear that the regression model has a better-quality metric.

On the other hand regarding the kernel functions in Mora's regression models [7], it is observed that the standard and min-max normalizers in that order have the best compromise in their quality metrics while in the present study the opposite occurs; min-max and standard.

The present study and that of Mora [7] achieve rbf as the best kernel function in the tuning, the rational quadratic function as the second best kernel function in Mora [7], and radial basic in the present study, all the results in the best normalizer. Regarding the second normalizer, it is observed that the alternative kernel functions gain prominence for both studies, with the min-max normalizer, the rational quadratic, and radial basic kernel functions standing out in Mora [7] as well as in the present study using the standard normalizer.

Finally, when visualizing the extrapolation and interpolation of data in Figure 4, it is observed that the irradiance supply in Nariño has similar segmentation patterns, with more discrepancy in the northern part of the department. However, since there is



less variability in the classification model, it is easier to identify the high, medium, and low irradiance zones.

## V. CONCLUSIONS

For obtaining classification models applying the SVM algorithm using the Landsat dataset, it is best to use a discretization by equal ranks with standard normalization and with the rbf kernel function, as it has the best compromise in accuracy and time as shown in Figure 3.

The rational quadratic and radial basic alternative kernel functions have a better compromise in accuracy than the rbf function using the max-min normalizer. However, the training and inference time of these functions is longer.

Using the normalizer, the Canberra and Triangular alternative kernel functions excel even in each discretizer, however, the accuracy values are lower than other configurations.

To discretize the irradiance, the best way to obtain irradiance classification models is to use the equal, uniform and kmeans discretizers in that order.

## AUTHORS' CONTRIBUTION

**Dalila-Mercedes Pachajoa:** Formal Analysis, Data Preprocessing, Research, Methodology, Software, Validation, Visualization, Writing-Original Draft, Writing-Revision and Editing.

**Héctor Mora-Paz:** Formal Analysis, Data Preprocessing, Research, Methodology, Software, Validation, Visualization, Writing-Original Draft, Writing-Revision and Editing.

**Dagoberto Mayorca-Torres:** Conceptualization, Methodology, Validation, Writing-Original Draft, Writing-Revision and Editing.

## REFERENCES

- [1] M. Zorzi, A. Chiuso, "The harmonic analysis of kernel functions" *Automatic*, vol. 94, pp. 125-137, 2018.  
<https://doi.org/10.1016/j.automatica.2018.04.015>
- [2] L. A. Belanche Muñoz, "Developments in kernel design", *UPCommons*, pp. 369-378, 2013.  
<https://upcommons.upc.edu/handle/2117/23278>



- [3] C. Spiegel, J. I. Cifuentes, *Definition and information on renewable energies*, Grade Thesis, Universidad de San Carlos de Guatemala, Guatemala, 2016
- [4] J. Gómez Ramírez, *Photovoltaic solar energy in Colombia: potentials, antecedents and perspectives*, Grade Thesis, Universidad Santo Tomás, Bogotá D.C., Colombia, 2018. <https://repository.usta.edu.co/handle/11634/10312>
- [5] Unidad de Planeación Minero Energética, *Integración de las energías renovables no convencionales en Colombia*, Bogotá D.C., Colombia, 2015
- [6] O. E. Cabrera Rosero, A. D. Pantoja Bucheli, " Analysis of the wind resource using R in non-interconnected zones (ZNI) of the department of Nariño (Colombia)," in *Latin American Conference on the Use of R in Research + Development*, 2018. <http://sedici.unlp.edu.ar/handle/10915/72585>
- [7] H. A. Mora-Paz, *Comparison of kernels on prediction of supply of alternative energy sources*, Master Thesis, Universidad Internacional de la Rioja, Logroño, Spain, 2021. <https://reunir.unir.net/handle/123456789/10020>
- [8] O. Cabrera, B. Champutiz, A. Calderon, A. Pantoja, "Landsat and MODIS satellite image processing for solar irradiance estimation in the department of Narino-Colombia," in *XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA)*, 2016, pp. 1-6. <https://doi.org/10.1109/STSIVA.2016.7743306>
- [9] S. H. Monger, E. R. Morgan, A. R. Dyreson, T. L. Acker, "Applying the kriging method to predicting irradiance variability at a potential PV power plant", *Renewable Energy*, vol. 86, pp. 602-610, 2016. <https://doi.org/10.1016/j.renene.2015.08.058>