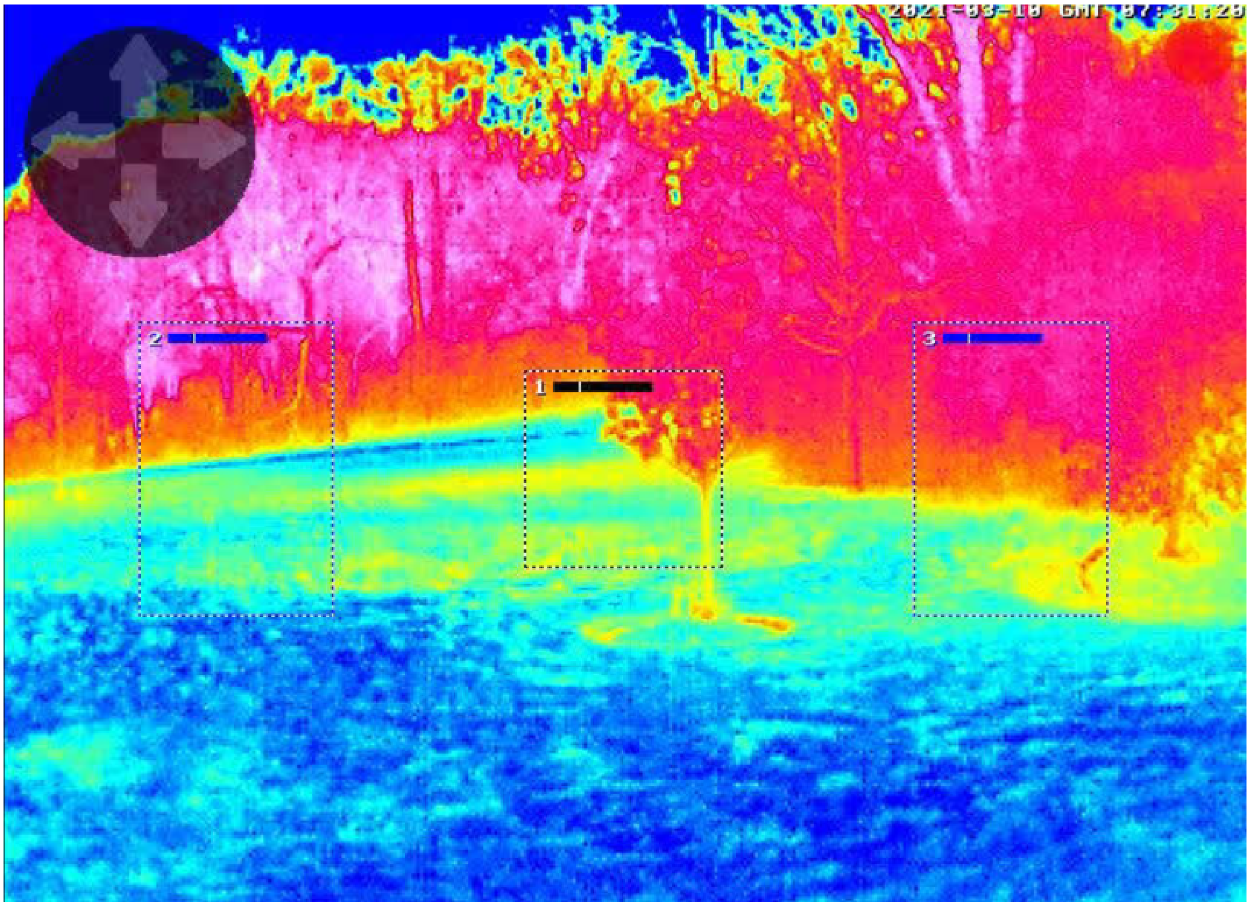
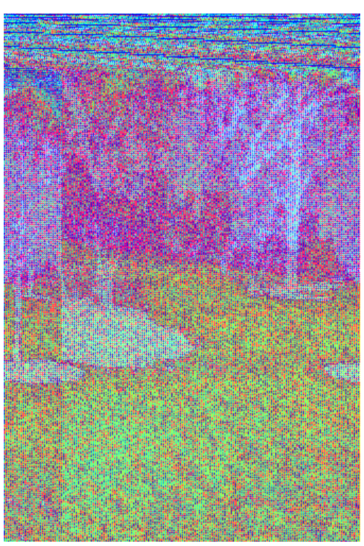
To begin, lifted condensation level, or the height at which an air parcel reaches 100\% humidity, is vital for the prediction and tracking of storms\cite{1}. Lifted condensation is often simply known as the base of a cloud. Rather than being calculated, lifted condensation has often been used in other machine learning models, including one for global horizontal irradiance[2]. The other focus of this study, solar irradiance, is extremely important for the efficient operation of solar power plants, and to predict/and or forecast it is crucial to setting up new plants[3]. With climate change becoming an ever bigger issue, and with solar power plants providing a cheap and reliable alternative, the calculation of solar irradiance becomes ever more important. With the advancement of machine learning technology, and potential for rapid calculation of these measurements without the need for specialized equipment, the potential utilization of both solar irradiance and lifted condensation level is great. The great benefit of models like the ones created are that the only code that needs to run is a simple input command into the model, which helps facilitate research, recording, and observation by people who otherwise may not be familiar with coding. While this is a novel approach, similar studies have been conducted on convection to predict storms over the Indian Ocean, using several different meteorological measurements, like humidity, wind speed, and pressure, and another used nowcasted Solar Irradiance within 15 minute to 1 hour intervals, using deep learning techniques to process several different images, as well as ground meteorological measurements[4, 5].Similar to this study, the previous studies used Scikit-learn modules and deep learning techniques to create different models.

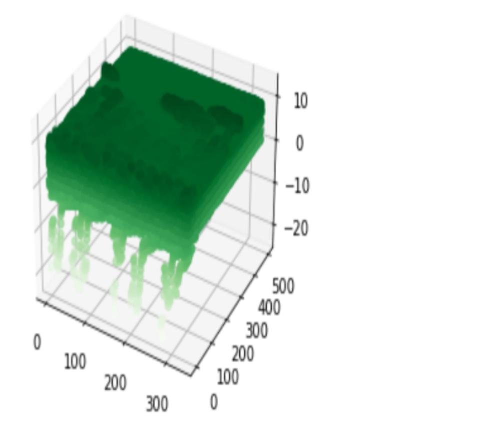
In particular, solar irradiance forecasting is a well-researched topic, and has had many potential avenues explored [5,6,7]. Most of these strategies heavily rely on images, both thermal and otherwise, in order to calculate the Solar Irradiance. In this study, rather than images, the arrays derived from the measurements the thermal cameras use to create the thermal images will be used to calculate the solar irradiance and lifted condensation level. The three main libraries for the execution of this are Tensorflow/Keras, Pytorch, and Scikit-learn. As mentioned before, Scikit-learn has been used in similar studies to predict similar values. Tensorflow/Keras has also been used to model climate and convection in other studies, and Pytorch has been used for image processing in ocean climate models [8,9].

Data

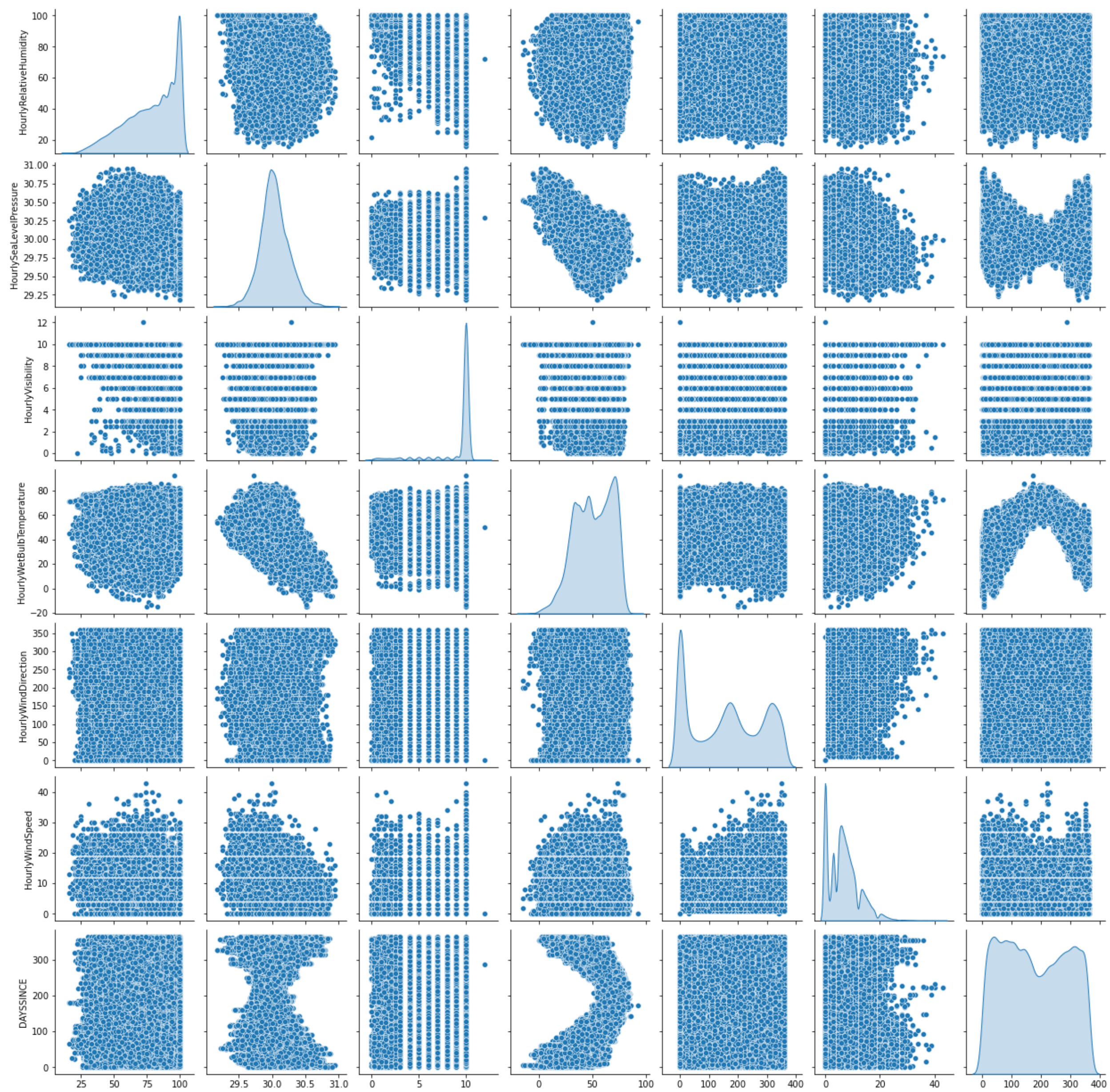


The first, and perhaps most crucial dataset used to train the model was a set of 1524 thermal images taken from 4/15/21 to 4/23/21 at Argonne National Laboratory. While the images themselves were largely unhelpful, each image came with a 2D temperature array of at least 252x336 values. These were paired with the other datasets in order to potentially predict other values.

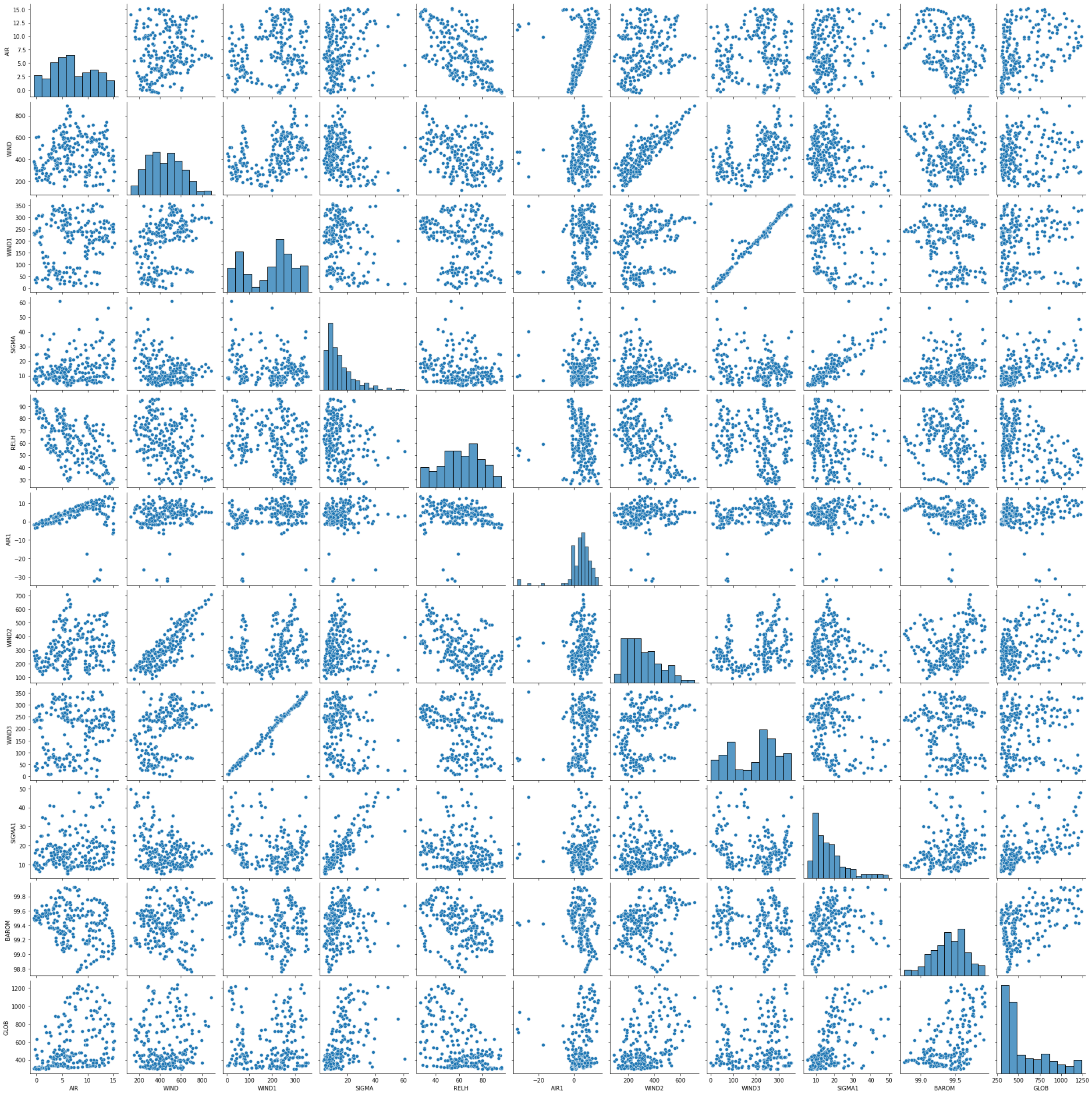




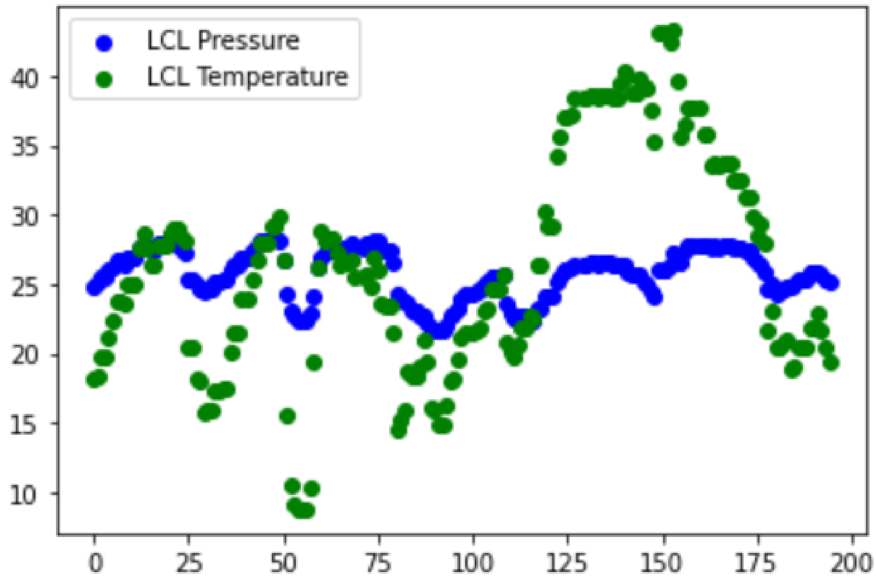
Above are two examples of visualizing the temperature arrays. Note the clear delineation of the trees, soil, and grass in the upper right. The temperature arrays are a clear representation of the thermal image, but as shown by the stretch and dimensions, they are not quite the same. This is once again clear on the right, which shows yet another representation of the temperature array. The x and z axes represent the position of the temperature values. The dropoff towards the left is a clear representation of the sky and yet another example of the utility and accuracy of the data collected by the thermal images.

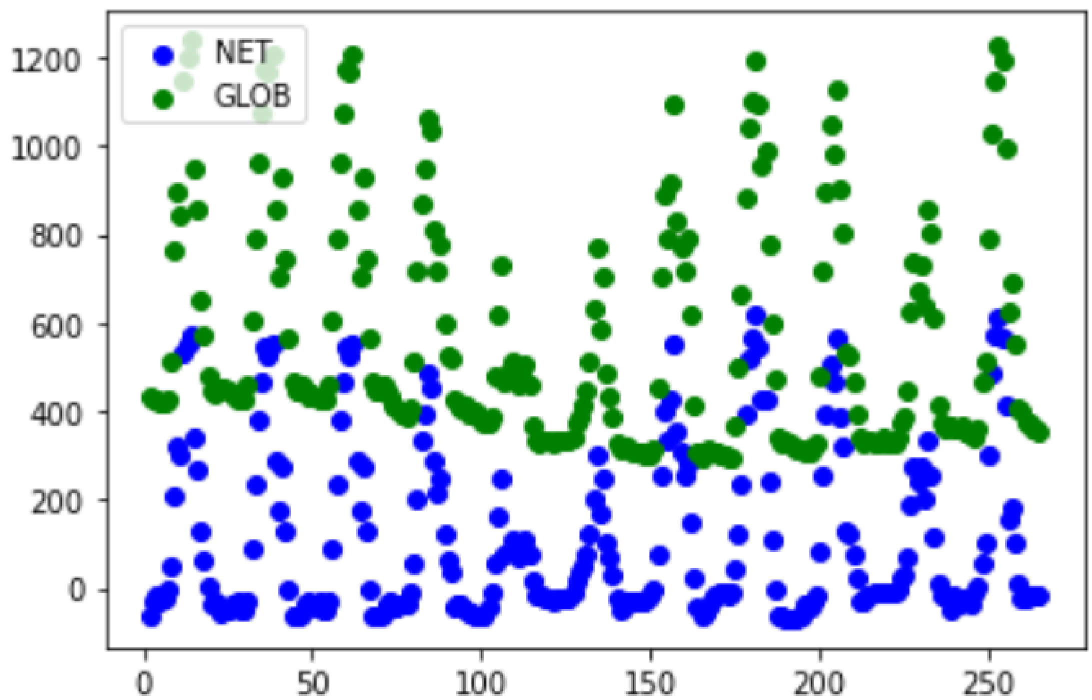


Several datasets were used to train the models. The first was the Local Climatological Dataset from the National Centers for Environmental Information. The dataset has data from across the US, but the data plotted above is from O'Hare Airport. Using the data collected, a second dataset of LCL pressures and temperatures(the pressures and temperatures at the LCL) were able to be calculated from the dry bulb temperature, the dew point temperature, and the station pressure. These features were then removed from the dataset, which was then turned into the above seaborn plot. Subsequently, the dataset was aligned with the thermal images, so that each thermal image was paired with a row from the Local Climatological Dataset that was recorded within an hour from when the thermal images were taken.



The next and final dataset is the pyrometer and associated data. Over the same period as the other datasets, a pyrometer was set up at Argonne National Laboratory, along with other weather equipment. This created a dataset with net and global irradiance, along with other features, displayed above. This was combined with the thermal images in much the same way the Local Climatological Dataset was.





These two graphs display the four values that the following models attempt to predict. As displayed, global irradiance and LCL temperature both retain a wide range, while LCL pressure and net irradiance retain a smaller, but more consistent spread of values.

Results

Library	Model Type		Prediction	Accuracy	MSE
Scikit-learn	SVR		LCL Temperature	0.196	1.675
Scikit-learn	SGDRegressor		LCL Temperature	-4.08E+34	1.021
Scikit-learn	Bayesian Ridge		LCL Temperature	0.34	1.52
Scikit-learn	LassoLars		LCL Temperature	1.11E-16	1.87
Scikit-learn	ARDRegression		LCL Temperature	0.33	1.53
Scikit-learn	PassiveAggressive Regressor		LCL Temperature	-13.68	7.67
Keras	**Sequential**		**LCL Temperature**	**0.98**	**0.304**
Keras	Sequential		LCL Pressure	0.755	0.8633
Keras	**Functional**		**LCL Pressure**	**0.977**	**0.576**
Keras	**Functional**		**Net Solar Irradiance**	**0.94**	**38.6**
Keras	Sequential		Net Solar Irradiance	0.997	9.77
Pytorch	NN		Net Solar Irradiance	-0.136	254.69
Keras	Sequential	Local Climatological Data	Global Solar Irradiance	0.95	54.167
Keras	Functional		Global Solar Irradiance	0.9019	60.53
Keras	**Sequential**	**Pyrometer**	**Global Solar Irradiance**	**0.925**	**83.055**

The best/most reliable models are highlighted above. Note that LCL pressure and net irradiance require the Functional models, which are the only ones(other than the Pytorch model) that require both the thermal image inputs and the corresponding datasets. It is curious that the values with less variance and more consistency are also more consistent with the thermal image, perhaps warranting further exploration

Conclusion

In summary, through machine learning, models to accurately calculate and predict the LCL pressure and temperature, as well as the global and net irradiance within reasonable limits have been created. The required models for each value differed greatly, with the LCL pressure and net irradiance requiring the thermal images for success, but the opposite being true for the LCL temperature and global irradiance. All in all, this study answers the question of possibility in the creation of a model that can determine irradiance and LCL from other meteorological data points, but also creates more problems- namely, the unexplained superior performance of certain models than others. Hopefully, this study and the models created will someday be utilized in the many potential applications of accurate prediction of lifted condensation level and solar irradiance.