Assessment and Correction of Solar Radiation Measurements with Simple Neural Networks

Jason Kelley

Department of Soil and Water Systems, University of Idaho; Moscow, ID 83844, USA; jasonrk@uidaho.edu; Tel.: +1-208-885-1015

Received: 16 July 2020; Accepted: 23 October 2020; Published: 27 October 2020

Abstract: Solar radiation received at the Earth’s surface provides the energy driving all micro-meteorological phenomena. Local solar radiation measurements are used to estimate energy mediated processes such as evapotranspiration (ET); this information is important in managing natural resources. However, the technical requirements to reliably measure solar radiation limits more extensive adoption of data-driven management. High-quality radiation sensors are expensive, delicate, and require skill to maintain. In contrast, low-cost sensors are widely available, but may lack long-term reliability and intra-sensor repeatability. As weather stations measure solar radiation and other parameters simultaneously, machine learning can be used to integrate various types of environmental data, identify periods of erroneous measurements, and estimate corrected values. We demonstrate two case studies in which we use neural networks (NN) to augment direct radiation measurements with data from co-located sensors, and generate radiation estimates with comparable accuracy to the data typically available from agro-meteorology networks. NN models that incorporated radiometer data reproduced measured radiation with an $R^2$ of 0.9–0.98, and RMSE less than 100 Wm$^{-2}$, while models using only weather parameters obtained $R^2$ less than 0.75 and RMSE greater than 140 Wm$^{-2}$. These cases show that a simple NN implementation can complement standard procedures for estimating solar radiation, create opportunities to measure radiation at low-cost, and foster adoption of data-driven management.

Keywords: solar radiation measurement; machine learning; data quality control; surface energy budget

1. Introduction

Reliable measurements of solar radiation are critical to accurately estimate the energy available for all other meteorological processes. As a result of water’s heat capacity and the latent heat of evaporation, a large portion of solar energy received at the surface is abstracted through evaporation from soil and water surfaces, and through transpiration from vegetation. Correspondingly, the flux of water from terrestrial surfaces is an important hydrologic linkage at local, regional, and global scales [1,2]. Energy-based methods to calculate evapotranspiration (ET) are central to regional and global water management, particularly for agriculture [3,4]. Downwelling shortwave radiation measurements from regional agro-meteorology networks are used to calculate reference ET, and combined with other weather data, are used determine crop water requirements via energy-based methods such as the Penman–Monteith equation [5–7]. Satellite images of surface temperatures and surface brightness are used to map spatial variability in ET and determine basin-scale water balances, and require corroborating surface measurements to calibrate [8–10]. Instantaneous and time-averaged measurements of net radiation (shortwave and longwave) are also used in calculating closure of the surface energy budget (SEB). The SEB is calculated from measurements of sensible heat flux, latent heat flux (evapotranspiration), and changes in energy storage (primarily in soil), with the sum of these terms...
equated with the measured net radiation. A closed SEB lends confidence to direct measurements of surface fluxes by methods such as eddy-covariance, with an expectation that the SEB residual does not exceed 10–20% of energy provided by net radiation [11–13]. Solar radiation models are also important boundary conditions that link climate and hydrologic models [14,15].

Despite the critical role of solar radiation in many environmental processes, direct measurements of solar radiation are sparse; and the data that are available sometimes lack adequate quality assurance. As a result, there is an abundance of empirical models that are used to approximate actual solar radiation, and global calculations which allow end users to assess the reliability of the resulting information. Typically, these empirical calculations begin with the calculation of exoatmospheric radiation, i.e., radiation received at the top of the atmosphere, based on latitude, day of year, and the solar constant [3,16]. To estimate the actual irradiance at the surface, the global value is adjusted downward, principally to account for cloudiness. Empirical studies have been conducted to correlate atmospheric transmissivity to more easily measured environmental parameters such as daily average temperature [17–19], weather conditions [20], or from regional parameters that can be estimated from synoptic scale weather, such as aerosol and precipitable water concentration [21]. Most of these approximations are valid over daily or longer timescales only. For meteorology networks that report reference ET for agriculture, measured downwelling shortwave radiation is typically used to estimate the actual irradiation. The shortwave irradiation is used in conjunction with a terrestrial longwave emission model to approximate the net radiation over shorter (hourly) time steps [16,22].

Increasingly, machine learning (ML) plays a role in prediction and modelling of environmental processes at global and regional scales. As early as twenty years ago, there was widespread interest in using various ML methods to tackle complex problems in hydrology and atmospheric science [23,24]; an ASCE special committee convened to survey hydrologic applications of neural networks [25,26]; and neural networks were compared to stochastic methods for predicting radiation [27]. Subsequent work that was specific to environmental applications has focused on input determination and calculation methods [28–31], data assimilation, down- or upscaling [32–34], and prediction of specific hydrologic and atmospheric processes such as precipitation and precipitable water [35,36], ground water flow [37], evapotranspiration [38,39], and plant phenology [40,41]. Recently, there are increasing numbers of methods presented which model time-varying process, including solar irradiation [42–45]. While far from comprehensive, these represent a substantial effort to develop ML methods that focus on prediction of environmental variables. In contrast, there are few published studies focusing on ML as a tool to evaluate site-specific processes and compare on-site environmental sensor data.

In this context, two season-long experiments were conducted in which solar radiation was measured to help evaluate SEB closure and estimate reference and actual ET. Both of these experiments primarily focused on the eddy-covariance method to determine ET and sensible heat flux over short (15 min) averaging periods, and in both cases, there were complications in measuring actual solar radiation. In the first experiment (referenced below as 2015), the net radiometer was fully operation for only the first half of the six-month experiment duration. One component (the uplooking pyranometer) failed midseason, leaving a critical gap in field measurements. Although downwelling shortwave radiation could be approximated by a number of methods as described above, none of these methods were appropriate over such short time periods as 15 min. The abundance of supplementary information, including the other three component radiometers, suggested the possibility of applying neural networks to estimate the actual, site-specific irradiance for fifteen-minute averaging periods. The primary goal of the second experiment (referred below as 2017) was to explore the ability of machine learning to estimate evapotranspiration by training a neural network to reproduce high-quality measurements from data obtained with low-cost sensors. As part of this experiment, a simple integrated net radiometer and other sensors were deployed for a three-month growing season, while a four-component radiometer was co-located for two short periods to train the neural network and validate the model output. ML methods were not employed to evaluate time-varying processes or for gap-filling of missing data, nor did they attempt to predict future conditions. Rather, the neural networks implemented.
pattern-recognition techniques to estimate instantaneous solar radiation received at the surface. The neural networks were trained by identifying characteristic relationships with other meteorological measurements such as air temperature, or in the case of downwelling shortwave radiation, using measurements of other radiation components (e.g., longwave and upwelling). In both cases, the expected outcome was a trained neural network that could reproduce other radiation estimation methods with higher fidelity (over time intervals shorter than one day) with less requirement for regional or other empirical calibrations, while utilizing in-situ environmental sensor measurements.

2. Experiments

Measurements of downwelling visible and photosynthetically active (PAR) solar radiation, net solar radiation (long- and shortwave), and other environmental data from two field experiments conducted in 2015 and 2017 are utilized in this analysis. These experiments were described previously, although these previous analyses were principally concerned with eddy-covariance flux measurements, and radiation data were not analyzed in any great detail previously.

For the first experiment, which was part of an effort to calibrate remote sensing measurements of ET, solar radiation and other meteorological sensors were installed with an eddy-covariance flux measurement system for a six-month period from 1 April to 1 October 2015 in an irrigated pasture near Rifle, Colorado, USA. A complete description of this field experiment can be found in a report published by the Upper Colorado River Commission [46]. Instrumentation included a four-component net radiometer (Model NR01, Hukseflux, Delft, The Netherlands), and two thermohygrometers (Model HMP155, Vaisala Oyj, Helsinki, Finland). Sensors were measured at five-second intervals, and average and standard deviations of measurements were recorded at 15 min intervals by a datalogger (Model CR1000, Campbell Scientific, Logan, UT, USA). Data from other sensors were not used in this analysis. During the course of this field experiment, the upward-looking shortwave radiometer (pyranometer) suffered an electronic failure and stopped reporting data (from this component sensor only) starting on 20 June 2015. Table 1 lists the sensors and periods of operation from which data used in this study were obtained. Downwelling solar radiation was also measured at the Rifle, CO Wildland Fire Remote Automated Weather Stations (RAWS) located 13 km NW of the experiment site. The 2015 data from this automated station (Station ID RFEC2) were obtained through the Mesowest data portal (https://mesowest.utah.edu/).

Table 1. Sensors from 2015 Field Experiment used in this analysis.

| Measured Parameter                      | Sensor (Manufacturer)                                      | Period of Operation  |
|-----------------------------------------|------------------------------------------------------------|----------------------|
| 4 component Net Radiation (excluding Downwelling SR) | NR01 (Hukseflux, Hukseflux, Delft, The Netherlands)         | 1 April to 1 October |
| Downwelling Shortwave Radiation         | NR01 (Hukseflux, Hukseflux, Delft, The Netherlands)         | 1 April to 20 June   |
| Air Temperature and Humidity (two heights) | HMP155 (Vaisala, Vanta, Finland)                           | 1 April to 1 October |

The second field experiment was conducted to evaluate real-time irrigation decision support, in which machine learning was used to estimate crop water demand. An eddy-covariance system and supplementary meteorological sensors were installed for 85 days (12 June to 4 September 2017) adjacent to a 120 ha irrigated field near Corvallis, Oregon, USA. In this analysis, data were included from a thermopile net radiometer (Model Q-7, Radiation and Energy Balance Systems, Inc., Seattle, WA, USA), a PAR quantum sensor (Apogee Instruments, Inc. Logan, UT, USA), and thermohygrometers (Model HMP60, Vaisala Oyj, Helsinki, Finland). For a 12-day period from 21 June to 3 July, a NR01 four-component radiometer and additional instrumentation were co-located with the other sensors. A complete description of this field experiment was provided in previous publications [47,48]. For a 12-day period—19–31 August—the NR01 radiometer and other sensors were deployed again at
the experimental site as part of a large coordinated experiment on atmospheric turbulence during the Great American Eclipse, as described previously [49]. The sensors and periods of operation from which data were used in this study are listed in Table 2.

Table 2. Sensors from 2017 Field Experiment used in this analysis.

| Measured Parameter                  | Sensor (Manufacturer)                                      | Period of Operation       |
|-------------------------------------|-----------------------------------------------------------|---------------------------|
| Net Radiation                       | Q-7 net radiometer (REBS, no longer in production)        | 12 June to 4 September    |
| Net Radiation                       | NR01 (Hukseflux, Hukseflux, Delft, The Netherlands)        | 21 June to 3 July         |
|                                     |                                                           | 19 to 31 August           |
| Air Temperature and Humidity        | HMP60 (Vaisala, Vanta, Finland)                            | 12 June to 4 September    |
| Photosynthetic Active Radiation     | PAR quantum sensor (Apogee Instr., Logan, UT, USA)         | 12 June to 4 September    |
| Wind Speed and Direction            | DS-2 Sonic anemometer (Decagon Inc., Pullman WA, USA)     | 12 June to 4 September    |
| Wind Speed and Direction            | Wind Sentry cup and vane (RM Young, Traverse City, MI, USA)| 12 June to 4 September    |

Experimental and meteorological network data were converted from raw text files and processed in Matlab software version 9.5 (The Mathworks Inc., Natick, MA, USA, 2018). Public data obtained for this analysis were checked for quality control flags, although no data were marked as suspect. For the analyses shown in the following results, no records were automatically removed from the experimental data to avoid biasing the subsequent analysis. Neural network methods in this analysis used Mathwork’s Machine Learning and Deep Learning Toolbox in Matlab. As implemented in the toolbox, neural networks (NNs) are structured programmatically in multiple layers of nodes. In each node, weighting and bias vectors and a transfer function control how input parameters are converted to nodal outputs. Environmental sensor data from the two field experiments, such as air temperature and humidity, are used as inputs parameters in the NN models. Consistent with the relatively small datasets used in this study, each layer was assigned the same number of nodes (2–5) as the number of input variables. The first node layer receives as input a number of records of all input variables, and each subsequent layer of nodes receives and applies weighting and bias functions in turn until all nodal outputs are combined to return a final model estimate. In the case of supervised learning, each record used in training is also associated with a control estimate of the output parameter to be modelled. For all the following NN models, the modelled output parameter was either downwelling shortwave radiation or spectrally-integrated net radiation. The control estimates used for training the NNs were measured values of the corresponding variable.

Network training occurs iteratively to adjust weights and biases for the nodes, and tested against a subset of the batched training records. For this study, all instances of training used 90% of the records for the training process (to estimate the output), and 10% assigned to the validation and testing phases which identifies the generality of the network model and prevents overfitting [50]. Supervision refers to the comparison of the control estimate to the NN output, and optimization involves minimizing network output error by adjusting the nodes. Here, mean square error (MSE) was the learning statistic used in backpropagation, in which a random subselection of the training data is compared to the corresponding model output. The backpropagation method is conducted repeatedly in epochs, re-initializing the random assignment of training records before proceeding with training in each epoch. The training proceeds by adjusting the weighting and bias vectors in order to minimize the difference between the control ET measurement and the value predicted by the NN. Training epochs are conducted until any one of a number of criteria are reached—a minimized
MSE, a maximal limit of epochs, or an insignificant change or increase in MSE in subsequent epochs. Each training procedure is randomly initialized, so that two networks trained on identical data can produce slightly different outputs. In data-limited conditions such as on-farm monitoring, we have previously proposed that variability in the model output can approximately describe the robustness and generality of the solution [51]. Simplified measures of model robustness (such as RMSE and $R^2$) are appropriate descriptors of model reliability when the model utilized simple input datasets—for example [39]—as opposed to models which enable multi-linear regression with several independent factors, as demonstrated in [41].

In both field experiments shown here, NNs were used to model solar radiation data. In the case of the 2015 experiment, the objective of modelling radiation was to synthesize the downwelling shortwave irradiation after one component of the net radiometer failed. NN models were trained on data collected over a three-week period from 20 June to 10 July prior to the failure of the uplooking pyranometer. Each model used a different configuration of inputs, including air temperature and humidity, uplooking pyrgeometer, and downlooking pyranometer data, and was trained to reproduce the (downwelling) shortwave irradiation measured by the uplooking pyranometer. The trained NNs then modelled shortwave irradiation for the entire 175 days of input data records, including during the 100 day period when the uplooking pyranometer ceased to operate.

For the 2017 field experiment, a four-component radiometer was co-located with other sensors for a 12 day period only (21 June to 3 July), and the objective was to estimate site-specific net radiation with a minimum set of sensors as or more accurately than the integrated thermopile radiometer, compared to a simplified net radiation model [7,22,52]. NNs were trained with various input configurations including data from the Q-7 net radiometer, a photosynthetically active radiation (PAR) sensor, air temperature/humidity, and wind speed sensors for all records during the 12-day period. The resulting NNs were then used to estimate net radiation over the entire 85-day experiment. Because the structure of these NNs treat each record as an independent instance and ignore time dependent features such as ordering, input data for all NNs included the timestamp as a fraction of day. The different arrangements of input variables for each trained neural network are described in Table 3.

### Table 3. Neural Network Input scheme and designations in text.

| 2015 Experiment (Rifle, CO, USA) | 2017 Experiment (Corvallis, OR, USA) |
|----------------------------------|-----------------------------------|
| NN4    | Time, Uplooking IR, Air Temperature, RH |
| NN3    | Time, Uplooking IR, Air Temperature |
| NN3c   | Time, Air Temperature, Relative Humidity |
| NN2    | Time, Air Temperature |
| NN2b   | Time, Uplooking IR |
| NN4b   | Time, Downloading SW, Up. IR, Air Temp. |
| NN5    | Time, PAR, Wind Speed, Air Temp., RH |
| NN3    | Time, PAR, RH |
| NN3b   | Time, Air Temperature, RH |
| NN2    | Time, PAR |
| NN2b   | Time, Air Temperature |

Table 3. Data type designations: Time = Recorded time of day; IR = Pyrgeometer; RH = relative humidity; SW = Pyranometer; PAR = Photosynthetic Active Radiation (flux); “Downlooking” indicates a sensor pointing toward the ground; “Uplooking” indicates skyward-pointing sensor.

Unlike previous demonstrations that used NNs to estimate evapotranspiration [48,51], the control estimates of irradiance (2015) and net radiation (2017) were not measured for the duration of these experiments, so that another basis was needed to evaluate the NN output. For the 2015 experiment,
additional solar radiation data obtained from a nearby RAWS weather station provided a local, secondary reference with hourly records. Although the distance between the two stations was only 13 km, the extreme topographic relief meant that the instantaneous records were often decoupled/delayed in time. Therefore, NN outputs were compared to experimental measurements (while the sensor was still operating) and to RAWS data, at both hourly intervals and as summed energy over 24 h periods (MJ·m⁻² received per day), without making any assumption of lagged correlations at other timescales. Daily data and model outputs were also compared to clear-sky shortwave radiation following Chapter 4 and Appendix H of ASCE Manual 70 [16]. The last test calculated a synthetic surface albedo as the ratio of downwelling shortwave radiation to measured upwelling shortwave radiation, and compared this synthetic albedo to the measured albedo (as determined in the period prior to failure of the net radiometer). The measured and synthetic albedo values were binned by time of day to show the daily development of surface reflectance corresponding to solar angle, angle of incidence, and other conditions affecting reflectance (such as wetness or ground cover).

For the 2017 experiment, a second net radiometer of a different type (integrated thermopile) also collected data for the entire duration of records, indicated throughout this text by the model name Q-7. Data from this second net radiometer were used as an independent basis for comparing against the NN output, which was trained to reproduce net radiation measured by the four-component radiometer. Additionally, the ASCE Manual 70 method was used to calculate the effective net radiation using the shortwave irradiation measured by the PAR sensor. Total daily irradiation estimates produced by each of the NN and the ASCE models were compared against that measured by the Q-7 net radiometer.

3. Results

Implementing the training phase and subsequent estimation of radiation is straightforward using the interface of the Matlab toolbox. The resulting estimate of shortwave irradiation (for the 2015 experiment) and net radiation (for the 2017 experiment) followed expected diurnal patterns, and reproduced the incident radiation on clear and cloudy days, as well as days in which sunlight was reduced because of wildfire smoke (such as 23 August, as shown in Figure 1). NN models were even able to reproduce the total solar eclipse observed on 21 August 2017, with varying degrees of fidelity (Figure 1).

![Figure 1. Time series of measured (Q-7) and neural network models of net radiation, 2017 experiment. Three example days show a range of sky conditions, including total eclipse on 21 Aug.](image-url)

While superficially similar across a broad range of sky conditions, the NN models are comparable across the range of observed output values as well, and confirmed by independent measurements. There are differences between the different NN implementations; in the example above, NN3b and NN2b are not able to reproduce transients induced by cloudy conditions (Figure 1, 23 August). The results from this limited study demonstrate that the robustness of any particular NN implementation depends on the specific input parameters used. Previous investigations in which NNs were applied to
interpret sensor data have highlighted the sensitivity of trained networks to the representativeness of conditions used for training [48]. However, due to the constraints imposed by the availability of observations, the same training period was used in all NN implementations, so that the robustness of the NN models were not evaluated with respect to the selected training periods.

3.1. Neural Network Results for 2015 Experiment

Among the different input configurations used from the 2015 data, only those which included a radiation measurement were able to reproduce the observed downwelling radiation with reasonable fidelity (Table 4). Those models included NN4 (using uplooking pyrgeometer data), NN4b (downlooking pyranometer and uplooking pyrgeometer), NN3 (uplooking pyrgeometer), and NN2b (uplooking pyrgeometer). The performance of the two NN models using only time of day, temperature, and/or relative humidity NN3c and NN2 was considerably worse (Figure 2d,e).

In order to evaluate model performance beyond the period when the NR01 net radiometer was operating, the albedo was calculated as the ratio of upwelling to downwelling shortwave radiation, and the median and quartiles of 15 min albedo values were binned by hour of day. The measured albedo was tabulated for the entire period that all four components of the NR01 operated. The synthetic albedo was tabulated for NN output records which did not include downlooking pyranometer data. Figure 3 shows that the NN4 synthetic albedo exhibits behavior more similar to the measured albedo, as compared to the NN2 model. The NN4 model uses uplooking pyrgeometer data, while the NN2 model did not include any radiometer data and exhibits greater scatter in albedo values at all times of the day.

| NN Model | R²     | RMSE (W·m⁻²) | Slope |
|----------|--------|--------------|-------|
| NN4      | 0.915  | 93.4         | 0.94  |
| NN4b     | 0.93   | 92.8         | 1.04  |
| NN3      | 0.896  | 100          | 0.905 |
| NN3c     | 0.737  | 148          | 0.763 |
| NN2      | 0.737  | 140          | 0.72  |
| NN2b     | 0.902  | 102          | 0.957 |

Table 4. NN models outputs compared to measured shortwave radiation. Data from training periods was excluded; samples included only data recorded on 1 April to 20 May, and 11 to 19 June.
Finally, the NN modelled irradiance time series were used to tabulate the daily energy received at the surface, and these data were compared to the daily energy reported at the nearby RAWS meteorological station (RFEC2). Figure 4 shows the time series of the six-month daily irradiance in MJ·m⁻²·d⁻¹ measured by the NR01, modelled by NN4 and NN4b, and reported at RFEC2. The data were plotted with the clear-sky solar radiation envelope (dashed line), which was calculated on a daily interval for the stations’ latitude and elevation.

The clear sky envelope indicates that both daily measurements may exceed physically realistic limits in April, and that RFEC2 station may also report physically unrealistic values in August and September. Consequently, the models trained on these values also exceed the clear sky radiation for some parts of the season. Nonetheless, the time series of the measured and modelled values generally correspond, as evident during the many cloudy days observed throughout May, and similar daily trends through July. Likewise, although daily irradiance (Rₚ) measured during the experiment appear to trend lower than those measured at RFEC2, the full range of output by the NN4 and NN4b generally corresponds to RFEC2 throughout the season (Figure 5).
Figure 3. Measured and “synthetic” albedo—i.e., the ratio of upwelling to downwelling shortwave radiation, binned by hour of day. All data inclusive of training periods. Red hash markers indicate a relatively small number of outlier values beyond the 1st and fourth quartiles.

Figure 4. Measured and modelled daily irradiance with the clear sky envelope (dashed line). Period of record used for training NN models is indicated by shaded box.
3.2. Neural Network Results for 2017 Experiment

As noted earlier, the NN models were able to reproduce the observed time series of net radiation under wide range of conditions, including a total solar eclipse which occurred during the 2017 experiment (the experiment site was located within the path of total solar eclipse). However, the accuracy of the five NN models did vary considerably, especially under these particularly unusual conditions (Figure 6). Model performance was evaluated under more normal operating conditions, utilizing a co-located Q-7 thermopile net radiometer for the duration of the experiment as an independent comparator.

Model results for the 2017 data are shown in Table 5, which lists the $R^2$, root mean square error (RMSE), and slope of the linear regression between all combinations of the two observations (Q-7 and NR01 sensors) and the various NN models. Although the training period for the 2017 experiment was shorter, the strong correspondence between the NN outputs and the Q-7 net radiometer measurements indicates the robustness of the method to approximate net radiation from many different combinations of input data, even those which did not include any radiation sensor data (NN3b and NN2b). The slope of the linear regression between the NR01 and the Q-7 sensors are opposite in direction for the training period and the eclipse experiment (Q-7 measured 90% of the NR01 in the earlier period, and measured 107% of NR01 values during the eclipse, Figure 7). This may be indicative of experimental error, as the Q-7 should have demonstrated a reduction in measured radiation due to degradation of the polyethylene wind shield; however, a post-season “sensor inversion” calibration did not indicate a significant difference between the uplooking or downlooking aspects of the Q-7.
Figure 6. Modelled shortwave irradiance on 21 August 2017, during the Great American Solar Eclipse.

Table 5. NN models outputs compared to measured net radiation. Included are comparisons of net radiation measured by Q7 and NR01 sensors during the training period and during the subsequent experiment conducted around the eclipse (ECL). Data from training periods were excluded.

| Predictor: Observed | R²   | RMSE (W·m⁻²) | Slope |
|---------------------|------|--------------|-------|
| NR01: Q7            | 0.980| 35.2         | 0.90  |
| ECL: Q7             | 0.977| 31.3         | 1.07  |
| Q7: NN5             | 0.974| 42.9         | 1.14  |
| ECL: NN5            | 0.967| 44.5         | 1.25  |
| Q7: NN3             | 0.982| 35.5         | 1.13  |
| ECL: NN3            | 0.952| 51.8         | 1.21  |
| Q7: NN3b            | 0.927| 74.6         | 1.15  |
| ECL: NN3b           | 0.921| 78.7         | 1.41  |
| Q7: NN2             | 0.988| 28.9         | 1.13  |
| ECL: NN2            | 0.971| 39.4         | 1.21  |
| Q7: NN2b            | 0.914| 78.9         | 1.11  |
| ECL: NN2b           | 0.925| 73.8         | 1.36  |
The corresponding analysis of the NN models demonstrate similar trends compared to the NR01 data, in contrast with comparison to Q-7 data (Figure 8). While the general trends indicate the robustness of the NN to reproduce observed five-minute records of net radiation, greater scatter is observed for higher flux magnitudes in estimates by NN3b and NN2b. Both of these NN models lacked radiation data as input, and only used temperature (NN2b) or temperature and humidity (NN3b). Since temperature and humidity (and the sensors used to measure these parameters) vary more slowly, the lack of radiation sensor inputs in these NN models is likely why they are unable to reproduce short-term transients (usually observed during cloudy sky conditions), but they are able to reproduce phenomena that vary slowly in time (such as diurnal patterns or the eclipse).
Figure 8. Correspondence between measured (x-axis) and NN model of net radiation. The black markers indicate the full-season record of comparison (with Q-7 measurements) and the red markers indicate the short period of record around the 2017 solar eclipse (NR01 measurements). Results using NN5 are shown in 8(a,b). 8(c–f) show results from NN3, NN3b, NN2, and NN2b, respectively, including comparisons against both Q-7 sensor (black) and NR01 (red markers).

In keeping with the intent to evaluate the ability of the model to estimate the energy available for ET and other surface fluxes, the modelled net radiation was compared to the net radiation estimations used by the ASCE and FAO56 standardized reference ET methods [16,22]. As part of the standardization, these methods use only the measured downwelling shortwave radiation, calculate a cloudiness function by comparison to the clear sky radiation, and applying a standard albedo of 0.23 (based on a standardized grass reference surface). Terrestrial longwave and longwave backscatter from the sky are estimated using empirical functions that employ the measured air temperature and relative humidity. The standardized net radiation is the sum of the calibrated downwelling shortwave minus the reflected shortwave minus the emitted terrestrial longwave plus the backscattered longwave. In this case, the standardized method was employed with an independent, calibrated sensor (the Apogee PAR sensor) to determine the five minute and daily ASCE/FAO56 standardized net radiation. The standardized net radiation is plotted with the measured Q-7 and the five NN outputs, summed for daily periods, in Figures 9 and 10.
Figure 9. Daily net radiation measured by Q-7 sensor, modelled by five NN, and calculated by ASCE/FAO56 method. Period of record used for training models is indicated by shaded box.

Figure 10. Daily net radiation measured by Q-7 sensor compared to the NN models and standardized net radiation.

In Figure 9, note that the models were trained during the period 21 June to 3 July, and the closeness of the measured and modelled time series during this period does not indicate good model performance. Disregarding that period, there is still excellent correspondence between the Q-7 net radiation measurements and the NN models that utilized the PAR sensor (NN5, NN3, and NN2), with $R^2$ exceeding 0.97 in all three models, and an RMSE of 42.9, 35.5, and 28.9 Wm$^{-2}$ in the NN5, NN3, and NN2 models, respectively. Later in the season, the two models which did not utilize the PAR sensor (NN3b and NN2b) overpredicted the net radiation, with $R^2$ less than 0.93, and RMSE of 74.6 and 78.9 Wm$^{-2}$ in the NN3b and NN2b models, respectively. This is likely due to field conditions (temperature, humidity, albedo, etc.) which changed over the course of the season, and late in the season, observed air temperature and humidity were different from those observed during the training period.
4. Discussion

While it is unsurprising that the results demonstrate that simple neural networks can successfully reproduce observed radiation measurements, there have been few attempts to use these methods to corroborate or bias-correct available data. Usually, the provisioning of radiation data from meteorological networks, especially those that serve agriculture or other applied purposes, rely on two costly and tenuous propositions. First, sensors and stations must be maintained and calibrated regularly to prevent the development of long-term bias. While maintenance and calibration of radiometers are standard practice for network weather stations and research equipment, farm operators will not typically be able to justify these costs or spare the required effort among other farm activities. Secondly, sensor outputs must be checked against standardized methods following regionally specific calibrations, but these calibrations may not reflect site-specific conditions. These site-specific environmental conditions certainly do affect measurements, but more importantly they affect the actual energy available at the surface which drives all microclimatic and biological processes. Ultimately, a practical constraint on provisioning high-quality data will always be the availability of skilled technical work, both for sensor maintenance and data quality assurance, and this is especially apparent with large agricultural weather networks [53–55].

These results are a limited demonstration of simple neural networks which were used to recognize simple patterns and relationships in the data. This analysis has treated data as time-invariant, and more sophisticated ML methods such as long-term short-term machine learning (LSTM) may offer much better results and the ability to predict future trends [43,45] in ways that this simple approach did not. Earlier work with simple neural networks [48,51] indicates that the robustness of an NN model depends on providing the network with training data that are representative of a full range of conditions, and this is an area in which these results could be expanded. There is an indication that some actual measurement of radiation substantively improves the reliability of the radiation model. While it may not be obviously helpful to model that which is already being measured, it is likely that this supplementary sensor information will be helpful even if the absolute value of downwelling radiation cannot be measured. This would suggest the possibility of using extremely low-cost LED sensors to obtain relative light level measurements, and utilizing these in lieu of more expensive calibrated sensors, given that a calibrated sensor was available to train the low-cost sensor ensemble. These results also indicate that existing networks can serve as an information “backbone” that serves a much denser network of low-cost sensors, leveraging the existing monitoring resources to enable many more site-specific measurements across the landscape, as has been accomplished with other aspects of weather monitoring [56–58].

It can hardly be expected that a simple machine learning method will replace standardized methods such as provided by ASCE or FAO, which represent many careers worth of engineering and experimental effort, as well as many years of practical experience in field measurement. However, there will always remain a large portion of the landscape and conditions under which standard methods cannot be applied naively or without major technical expertise and effort. One simple, site-specific characteristic which dramatically affects energy availability is topographic aspect. Despite the fact that most areas are not perfectly flat, the technical considerations of estimating radiation on sloping surfaces [59] preclude this from the most simple applied methods. Even from this one simple example, the heterogeneity of the landscape should motivate us to adopt tools such as machine learning that can leverage all available data. In this way, ML can offer new insights, not by replacing existing methods but by augmenting our ability to recognize patterns in the landscape and by extrapolating our highest quality measurements to a broader range of conditions and places.

5. Conclusions

The application of simple neural networks to estimate solar radiation measurements was demonstrated by two case studies in which machine learning could complement standard estimating methods. In the case of a field experiment conducted in 2015, neural networks accurately estimated
missing data from a broken uplooking pyranometer, arguably the most important part of a four-component net radiometer. The replacement data were validated against the uplooking pyranometer data that were collected prior to the sensor failure, and also compared to a nearby RAWS meteorological station. The models that incorporated radiometer data (downwelling longwave and/or upwelling long- and shortwave) were able to reproduce the observed downwelling shortwave radiation with an $R^2$ greater than 0.9, and an RMSE less than 100 Wm$^{-2}$. Models that only relied on other weather parameters (such as air temperature or humidity) did not perform as well, with an $R^2$ less than 0.75, and an RMSE greater than 140 Wm$^{-2}$.

For the second case in 2017, neural networks were demonstrated in a field experiment, and used to estimate site-specific evapotranspiration with low-cost sensors. The field experiment provided data for machine learning by co-locating higher quality sensors with lower-cost sensors typically used in on-farm monitoring and agro-meteorological networks. This experiment was also able to leverage data from simultaneous experiments conducted during a total solar eclipse that occurred at the site. Neural networks were trained and tested using a variety of input data configurations. All models that utilized some form of radiation data (downlooking pyrgeometers or PAR sensors) were able to reproduce downwelling shortwave radiation or net radiation with high fidelity over five to fifteen minute measurement periods, with $R^2$ ranging from 0.97 to 0.98, and RMSE values less than 40 Wm$^{-2}$, as compared to observed values. Neural networks that only used air temperature, humidity, and timestamps estimated solar radiation measurements with less fidelity compared to observations, with $R^2$ ranging from 0.91 to 0.92, and RMSE values greater than 70 Wm$^{-2}$. These models also appeared to exhibit a stronger dependence on seasonal variation in other environmental conditions such as humidity and air temperature. Nonetheless, even the less accurate models were also able to approximate observed trends in measured radiation in many normal cases, and most models were able to reproduce the atypical and extreme drop in radiation observed during a total solar eclipse. These simple case studies suggest that further work should explore the range of valid conditions under which solar radiation could be modelled. Such neural network models would facilitate using low-cost sensors that have been trained, rather than calibrated, to report site-specific solar radiation.

**Funding:** This research was funded in part by USDA-AFRI under Grant No. 2017-67012-26125.

**Acknowledgments:** The author would like to thank Chad Higgins, who assisted with field work and provided equipment for the original research conducted in 2015 and 2017 and led the eclipse experiments.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**References**

1. Hegerl, G.C.; Black, E.; Allan, R.P.; Ingram, W.J.; Polson, D.; Trenberth, K.E.; Chadwick, R.S.; Arkin, P.A.; Sarojini, B.B.; Becker, A.; et al. Challenges in Quantifying Changes in the Global Water Cycle. *Bull. Am. Meteorol. Soc.* **2015**, *96*, 1097–1115. [CrossRef]

2. Fisher, J.B.; Melton, F.; Middleton, E.; Hain, C.; Anderson, M.; Allen, R.; McCabe, M.F.; Hook, S.; Baldocchi, D.; Townsend, P.A.; et al. The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources: The Future of evapotranspiration. *Water Resour. Res.* **2017**, *53*, 2618–2626. [CrossRef]

3. Irmak, S.; Irmak, A.; Allen, R.G.; Jones, J.W. Solar and net radiation-based equations to estimate reference evapotranspiration in humid climates. *J. Irrig. Drain. Eng.* **2003**, *129*, 336–347. [CrossRef]

4. Blonquist, J.M., Jr.; Bugbee, B. Solar, Net, and Photosynthetic Radiation. *Agroclimatol. Link. Agric. Clim.* **2020**, *60*, 1–49.

5. Gavilán, P.; Berengena, J.; Allen, R.G. Measuring versus estimating net radiation and soil heat flux: Impact on Penman–Monteith reference ET estimates in semiarid regions. *Agric. Water Manag.* **2007**, *89*, 275–286. [CrossRef]
6. Stöckle, C.O.; Kjelgaard, J.; Belloccchi, G. Evaluation of estimated weather data for calculating Penman-Monteith reference crop evapotranspiration. *Irrig. Sci.* 2004, 23, 39–46. [CrossRef]

7. Blonquist, J.M.; Allen, R.G.; Bugbee, B. An evaluation of the net radiation sub-model in the ASCE standardized reference evapotranspiration equation: Implications for evapotranspiration prediction. *Agric. Water Manag.* 2010, 97, 1026–1038. [CrossRef]

8. Chavez, J.L.; Gowda, P.H.; Howell, T.A.; Copeland, K.S. Radiometric surface temperature calibration affects on satellite based evapotranspiration estimation. *Int. J. Remote Sens.* 2009, 30, 2337–2354. [CrossRef]

9. Allen, R.G.; Irmak, A.; Trezza, R.; Hendrickx, J.M.; Bastiaanssen, W.; Kjaersgaard, J. Satellite-based ET estimation in agriculture using SEBAL and METRIC. *Hydrol. Process.* 2011, 25, 4011–4027. [CrossRef]

10. Morton, C.G.; Huntington, J.L.; Pohll, G.M.; Allen, R.G.; McGwire, K.C.; Bassett, S.D. Assessing calibration uncertainty and automation for estimating evapotranspiration from agricultural areas using METRIC. *JAWRA J. Am. Water Resour. Assoc.* 2013, 49, 549–562. [CrossRef]

11. Foken, T. The energy balance closure problem: An overview. *Ecol. Appl.* 2008, 18, 1351–1367. [CrossRef]

12. Leuning, R.; van Gorsel, E.; Massman, W.J.; Isaac, P.R. Reflections on the surface energy balance problem. *Agric. For. Meteorol.* 2012, 156, 65–74. [CrossRef]

13. Grachev, A.A.; Fairall, C.W.; Blonquist, B.W.; Fernando, H.J.S.; Leo, L.S.; Otárola-Bustos, S.F.; Wilczak, J.M.; McCaffrey, K.L. On the surface energy balance closure at different temporal scales. *Agric. For. Meteorol.* 2021, 281, 107823. [CrossRef]

14. von Schuckmann, K.; Palmer, M.D.; Trenberth, K.E.; Cazenave, A.; Chambers, D.; Champollion, N.; Hansen, J.; Josey, S.A.; Loeb, N.; Mathieu, P.-P.; et al. An imperative to monitor Earth’s energy imbalance. *Nat. Clim. Chang.* 2016, 6, 138–144. [CrossRef]

15. Trenberth, K.E.; Anthes, R.A.; Belward, A.; Brown, O.B.; Habermann, T.; Karl, T.R.; Running, S.; Ryan, B.; Tanner, M.; Wielicki, B. Challenges of a sustained climate observing system. In *Climate Science for Serving Society*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 13–50. ISBN 94-007-6691-2.

16. ASCE Manuals and Reports on Engineering Practice No.70: Evaporation, Evapotranspiration, and Irrigation Water Requirements, 2nd ed.; Jensen, M.E.; Allen, R.G. (Eds.) American Society of Civil Engineers: Reston, VA, USA, 2016; ISBN 978-0-7844-1405-7.

17. Allen, R.G. Self-calibrating method for estimating solar radiation from air temperature. *J. Hydrol. Eng.* 1997, 2, 56–67. [CrossRef]

18. Samani, Z. Estimating solar radiation and evapotranspiration using minimum climatological data. *J. Irrig. Drain. Eng.* 2000, 126, 265–267. [CrossRef]

19. Samani, Z.; Hargreaves, G.; Tran, V.; Bawazir, S. Estimating solar radiation from temperature with spatial and temporal calibration. *J. Irrig. Drain. Eng.* 2011, 137, 692–696. [CrossRef]

20. Thornton, P.E.; Running, S.W. An improved algorithm for estimating incident daily solar radiation from satellite data. *Ecol. Appl.* 1999, 93, 211–228. [CrossRef]

21. Majumdar, N.C.; Mathur, B.L.; Kaushik, S.B. Prediction of direct solar radiation for low atmospheric turbidity. *Sol. Energy* 1972, 13, 383–394. [CrossRef]

22. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. *Crop Evapotranspiration-Guidelines for Computing Crop Water Requirements-FAO Irrigation and Drainage Paper 56*; FAO: Rome, Italy, 1998; Volume 300.

23. Maier, H.R.; Dandy, G.C. Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications. *Environ. Model. Softw.* 2000, 15, 101–124. [CrossRef]

24. Gardner, M.W.; Dorling, S.R. Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmos. Environ.* 1998, 32, 2627–2636. [CrossRef]

25. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. Artificial neural networks in hydrology. I: Preliminary concepts. *J. Hydrol. Eng.* 2000, 5, 115–123. [CrossRef]

26. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. Artificial neural networks in hydrology. II: Hydrologic applications. *J. Hydrol. Eng.* 2000, 5, 124–137. [CrossRef]

27. Morid, S.; Gosain, A.K.; Keshari, A.K. Solar radiation estimation using temperature-based, stochastic and artificial neural networks approaches. *Hydrol. Res.* 2002, 33, 291–304. [CrossRef]

28. Bowden, G.J.; Maier, H.R.; Dandy, G.C. Input determination for neural network models in water resources applications. Part 2. Case study: Forecasting salinity in a river. *J. Hydrol.* 2005, 301, 93–107. [CrossRef]

29. Martí, P.; González-Altozano, P.; Gasque, M. Reference evapotranspiration estimation without local climatic data. *Irrig. Sci.* 2011, 29, 479–495. [CrossRef]
30. Kişi, Ö.; Öztürk, Ö. Adaptive neurofuzzy computing technique for evapotranspiration estimation. *J. Irrig. Drain. Eng.* 2007, 133, 368–379. [CrossRef]

31. García-Hinde, O.; Gómez-Verdejo, V.; Martínez-Ramón, M.; Casanova-Mateo, C.; Sanz-Justo, J.; Jiménez-Fernández, S.; Salcedo-Sanz, S. Feature selection in solar radiation prediction using bootstrapped SVRs. In Proceedings of the 2016 IEEE Congress on Evolutionary Computation (CEC), Vancouver, BC, Canada, 24–29 July 2016; pp. 3638–3645.

32. Irmak, A.; Kamble, B. Evapotranspiration data assimilation with genetic algorithms and SWAP model for on-demand irrigation. *Irrig. Sci.* 2009, 28, 101–112. [CrossRef]

33. Hunt, B.R.; Kostelich, E.J.; Szunyogh, I. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Phys. D Nonlinear Phenom.* 2007, 230, 112–126. [CrossRef]

34. Sachindra, D.A.; Ahmed, K.; Rashid, M.M.; Shahid, S.; Perera, B.J.C. Statistical downscaling of precipitation using machine learning techniques. *Atmos. Res.* 2018, 212, 240–258. [CrossRef]

35. Şenkol, O.; Yıldız, B.Y.; Şahin, M.; Pestemalcı, V. Precipitable water modelling using artificial neural network in Çukurova region. *Environ. Monit. Assess.* 2011, 184, 141–147. [CrossRef]

36. Najafi, M.R.; Moradkhani, H.; Wherry, S.A. Statistical downscaling of precipitation using machine learning with optimal predictor selection. *J. Hydrol. Eng.* 2011, 16, 650–664. [CrossRef]

37. Kourakos, G.; Harter, T. Vectorized simulation of groundwater flow and streamline transport. *Environ. Model. Softw.* 2014, 52, 207–221. [CrossRef]

38. Kumar, M.; Raghuvanshi, N.S.; Singh, R. Artificial neural networks approach in evapotranspiration modeling: A review. *Irrig. Sci.* 2011, 29, 11–25. [CrossRef]

39. Kisi, O. Least squares support vector machine for modeling daily reference evapotranspiration. *Irrig. Sci.* 2013, 31, 611–619. [CrossRef]

40. Dai, W.; Jin, H.; Zhang, Y.; Liu, T.; Zhou, Z. Detecting temporal changes in the temperature sensitivity of spring phenology with global warming: Application of machine learning in phenological model. *Agric. For. Meteorol.* 2019, 279, 107702. [CrossRef]

41. Meyers, J.N.; Kisekka, I.; Upadhyaya, S.K.; Michelon, G.K. Development of an Artificial Neural Network Approach for Predicting Plant Water Status in Almonds. *Trans. ASABE* 2019, 62, 19–32. [CrossRef]

42. Salcedo-Sanz, S.; Casanova-Mateo, C.; Muñoz-Mari, J.; Camps-Valls, G. Prediction of Daily Global Solar Irradiation Using Temporal Gaussian Processes. *IEEE Geosci. Remote Sens. Lett.* 2014, 11, 1936–1940. [CrossRef]

43. Qing, X.; Niou, Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy* 2018, 148, 461–468. [CrossRef]

44. Huang, C.-J.; Kuo, P.-H. A Deep CNN-LSTM Model for Particulate Matter (PM2.5) Forecasting in Smart Cities. *Sensors* 2018, 18, 2220. [CrossRef] [PubMed]

45. Liu, H.; Mi, X.; Li, Y. Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. *Energy Convers. Manag.* 2018, 159, 54–64. [CrossRef]

46. Cuenca, R.H.; Higgins, C.W. Assessing Agricultural Consumptive Use in the Upper Colorado River Basin Phase II; Upper Colorado River Commission. 2016, p. 326. Available online: http://www.ucrcommission.com (accessed on 9 July 2020).

47. Kelley, J.; Pardyjak, E.R. Using Neural Networks to Estimate Site-Specific Crop Evapotranspiration with Low-Cost Sensors. *Agronomy* 2019, 9, 108. [CrossRef]

48. Higgins, C.W.; Drake, S.A.; Kelley, J.; Oldroyd, H.J.; Jensen, D.D.; Wharton, S. Ensemble-Averaging Resolves Rapid Atmospheric Response to the 2017 Total Solar Eclipse. *Front. Earth Sci.* 2019, 7, 198. [CrossRef]

49. Beale, M.; Hagan, M.; Demuth, H. *Neural Network Toolbox (TM) Reference*; 9.1.; The MathWorks Inc.: Natick, MA, USA, 2016.

50. Kelley, J.; McCauley, D.; Alexander, A.; Gray, W.; Siegfried, R.; Oldroyd, H.J. Using Machine Learning to Integrate On-Farm Sensors and Agro-Meteorology Networks into Site-Specific Decision Support. *Trans. ASABE* 2020, 63. [CrossRef]
52. Allen, R.G.; Jensen, M.E. ASCE Manual 70–Second Edition: Evaporation, Evapotranspiration and Irrigation Requirements. In Proceedings of the 2015 ASABE/IA Irrigation Symposium: Emerging Technologies for Sustainable Irrigation–A Tribute to the Career of Terry Howell, Sr. Conference Proceedings; American Society of Agricultural and Biological Engineers, Long Beach, CA, USA, 9–11 November 2015; pp. 1–16.

53. Haigh, T.; Koundinya, V.; Hart, C.; Klink, J.; Lemos, M.; Mase, A.S.; Prokopy, L.; Singh, A.; Todey, D.; Widhalm, M. Provision of climate services for agriculture: Public and private pathways to farm decision-making. Bull. Am. Meteorol. Soc. 2018, 99, 1781–1790. [CrossRef]

54. Laurenson, M.R.; Kiura, T.; Ninomiya, S. Providing agricultural models with mediated access to heterogeneous weather databases. Appl. Eng. Agric. 2002, 18, 617. [CrossRef]

55. Laurenson, M.R.; Kiura, T.; Ninomiya, S. Providing agricultural models with mediated access to heterogeneous weather databases. Appl. Eng. Agric. 2002, 18, 617. [CrossRef]

56. Allen, R.G.; Pereira, L.S.; Howell, T.A.; Jensen, M.E. Evapotranspiration information reporting: I. Factors governing measurement accuracy. Agric. Water Manag. 2011, 98, 899–920. [CrossRef]

57. Mahmood, R.; Boyles, R.; Brinson, K.; Fiebrich, C.; Foster, S.; Hubbard, K.; Robinson, D.; Andresen, J.; Leathers, D. Mesonets: Mesoscale weather and climate observations for the United States. Bull. Am. Meteorol. Soc. 2002, 18, 617. [CrossRef]

58. Reges, H.W.; Doesken, N.; Turner, J.; Newman, N.; Bergantino, A.; Schwalbe, Z. COCORAHS: The evolution and accomplishments of a volunteer rain gauge network. Bull. Am. Meteorol. Soc. 2016, 97(10), 1831–1846. [CrossRef]

59. van de Giesen, N.; Hut, R.; Selker, J. The Trans-African Hydro-Meteorological Observatory (TAHMO). Wiley Interdiscip. Rev. Water 2014, 1, 341–348. [CrossRef]

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

© 2020 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).