



**HAL**  
open science

## **Evaluation of an Artificial Neural Network based method for surface heat flux estimates over heterogeneous landscapes**

Mathilde Jomé, Fabienne Lohou, Marie Lothon, Lilan Fautrez, Antoine Vial, Solène Derrien, Antoine Royer, Maurin Zouzoua, Eric Paradyjak, Jason Kelley, et al.

### ► **To cite this version:**

Mathilde Jomé, Fabienne Lohou, Marie Lothon, Lilan Fautrez, Antoine Vial, et al.. Evaluation of an Artificial Neural Network based method for surface heat flux estimates over heterogeneous landscapes. *Journal of the European Meteorological Society*, 2025, 3, pp.100021. <10.1016/j.jemets.2025.100021>. <hal-05349308>

**HAL Id: hal-05349308**

**<https://hal.science/hal-05349308v1>**

Submitted on 5 Nov 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License



Contents lists available at ScienceDirect

## Journal of the European Meteorological Society

journal homepage: <https://www.sciencedirect.com/journal/journal-of-the-european-meteorological-society>

## Evaluation of an Artificial Neural Network based method for surface heat flux estimates over heterogeneous landscapes

Mathilde Jomé<sup>a</sup>, Fabienne Lohou<sup>id a,\*</sup>, Marie Lothon<sup>a</sup>, Lilan Fautrez<sup>a</sup>, Antoine Vial<sup>a</sup>, Solène Derrien<sup>id a</sup>, Antoine Royer<sup>id a</sup>, Maurin Zouzoua<sup>id b</sup>, Eric Pardyjak<sup>c</sup>, Jason Kelley<sup>id d</sup>, Sophie Bastin<sup>id b</sup>

<sup>a</sup> LAERO, Université de Toulouse, CNRS, UT3, IRD, Toulouse, 31400, France<sup>b</sup> LATMOS/IPSL, Université Paris-Saclay, Sorbonne Université, CNRS, Guyancourt, 78280, France<sup>c</sup> Department of Mechanical Engineering, University of Utah, Salt Lake City, UT 84112, USA<sup>d</sup> Soil and Water Department, University of Idaho, Moscow, ID 83844, USA

## ARTICLE INFO

## Keywords:

Artificial Neural Network

Eddy-Covariance

Surface Turbulent Heat Fluxes

Surface Heterogeneity

## ABSTRACT

A new Artificial Neural Network (ANN)-based method for measuring surface-layer fluxes at low cost without compromising accuracy is presented. It is based on a one-year dataset of surface sensible and latent heat fluxes from eddy-covariance stations, along with classical meteorological variables from automatic weather stations, acquired over a meadow and a maize crop field. For each of these landscapes, ANNs are trained to estimate observed heat fluxes based on relevant variables that basically describe radiative and atmospheric forcing. The costly eddy-covariance station is required only for a short measurement period to collect sufficient reference data to train ANNs. Then, the estimation of heat fluxes relies on measurements from the affordable automatic weather station. The final objective is to deploy this method during a one-year field campaign to monitor different land-surface covers. Three questions are investigated here: the experimental scenario that allows to constitute the best training dataset, the accuracy of the ANNs estimates across different timescales, and the ANNs performance according to the land-surface cover. From a general perspective, the ANNs estimates have quite similar performance for all training datasets, irrespective of the vegetation cover. However, for conditions that were not well sampled in the training dataset, higher discrepancies were observed between ANNs estimated flux and those measured using eddy-covariance station. The comparison between the two vegetation covers indicated a degradation in ANNs performance for the maize crop compared to the meadow, with an increase of the bias up to 40 % and a decrease of the Pearson correlation coefficient up to 8 %.

## 1. Introduction

The measurement and understanding of surface-atmosphere exchanges are essential to study the evolution of the atmosphere in response to the surface at different space and time scales. These measurements control the surface energy balance and are therefore needed to assess and validate numerical weather prediction and climatic models (Boone et al., 2009; Lohou et al., 2014; Couvreux et al., 2016). When considering complex environments, documenting surface-atmosphere exchanges is even more challenging (Bou-Zeid et al., 2020; Shen et al., 2016). Indeed, heterogeneous landscapes are composed of many types of land covers (e.g., crops, meadow, forest, water, urban, etc.), each of them with a specific interaction with the atmosphere depending on its surface temperature and humidity, vegetation or cover character-

istics (e.g., leaf area index, albedo, roughness, etc.). Besides the type of surface and soil conditions, surface-atmosphere exchanges depend strongly on the meteorological conditions. Soil, vegetation and meteorology have both strong diurnal and seasonal cycles. Therefore, to efficiently estimate surface-atmosphere exchanges within heterogeneous numerical grid cells, the deployment of a network of measurement stations over different land cover types is required.

The interactions between the surface and the atmosphere were first measured and theorized under the assumption of a homogeneous surface and stationary flow. Monin-Obukhov similarity theory (MOST) was formulated tens of years ago (Foken et al., 2006) to describe fluxes of momentum, heat and other quantities in the turbulent surface layer, over homogeneous surfaces and with stationary flow. MOST predicts the flux-gradient relationship in terms of the measurement height and

\* Corresponding author.

E-mail address: [fabienne.lohou@utoulouse.fr](mailto:fabienne.lohou@utoulouse.fr) (F. Lohou).<https://doi.org/10.1016/j.jemets.2025.100021>

Received 25 November 2024; Received in revised form 1 October 2025; Accepted 11 October 2025

Available online 2 November 2025

2950-6301/© 2025 The Authors. Published by Elsevier B.V. on behalf of European Meteorological Society. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

the Obukhov length through empirical stability functions. However, the Earth's surface is often heterogeneous and there is no theory to describe the turbulence and flux-gradient relationship under such conditions (Nakamura and Mahrt, 2001). Uncertainties in some of the MOST universal functions have already been pointed out and may be associated with the effects of a heterogeneous surface or with important entrainment effects at the top of the boundary layer (Detto et al., 2008; Lohou et al., 2010; van de Boer et al., 2014). As noted by Moene et al. (2006), MOST is accurate as long as the dominant driving force is included in the scale that is used to normalize the scalar statistics.

Eddy-covariance stations are used to locally estimate and monitor land-atmosphere interactions on an occasional (i.e., measurement campaigns) or permanent basis (i.e., permanent reference station). The eddy-covariance method (Foken et al., 2012) is based on the high-frequency covariance of the vertical wind and air temperature (for the sensible heat flux) or water vapor mixing ratio (for the latent heat flux). Expensive fast response sensors are thus necessary. Moreover, an heterogeneous landscape composed of five to six main surface covers would need as many eddy-covariance stations, quickly becoming exceedingly expensive.

Recent studies (Kumar et al., 2002, 2010; Kelley and Pardyjak, 2019; Kelley, 2020; Kelley et al., 2020; Zouzoua et al., 2024) have proposed and tested a new way of estimating surface turbulent heat fluxes using Artificial Neural Networks (ANNs). The methodology is briefly as follows: the ANNs are trained to estimate eddy-covariance fluxes (observed surface sensible and latent heat fluxes) using basic near-surface meteorological variables as inputs. Kelley and Pardyjak (2019) used ANNs to estimate water requirements for two crops (vegetable beans and hazelnuts). Indeed, the ANNs were trained to estimate the actual evapotranspiration with measurements from affordable automatic weather stations. Conclusive results were obtained using just four physical parameters (near-surface temperature, humidity, wind speed and incoming solar radiation) and the time of day over a short training period of two weeks to estimate the evapotranspiration during two months. Kelley and Pardyjak (2019) conclude that extending the estimate of the evapotranspiration throughout the growing season would require a training set that takes into account the variability of meteorological conditions as well as the different growing stages of the vegetation. These results highlight the ease with which such a method can be applied while also underscoring the importance and difficulty of sampling enough variability (meteorological conditions, development of the vegetation, soil conditions, etc.). Zouzoua et al. (2024) used a data-driven statistical model trained on observational data to evaluate surface turbulent heat fluxes estimation in numerical models. They used a total of 13 variables to build their ANNs including net radiative flux, meteorological conditions, soil moisture and temporal coordinates. However, three variables showed the strongest influence on estimated fluxes: the net radiative flux, the mean potential temperature and soil moisture. Kelley and Pardyjak (2019) found the contrary concerning this last variable. In their case, adding information related to the soil moisture deteriorates the ANNs performance. However, the conditions and objectives of the two studies are quite different, and while Zouzoua et al. (2024) trained their ANNs with several years of observational data, Kelley and Pardyjak (2019) only used a few months. Several studies (Kelley and Pardyjak, 2019; Kelley, 2020; Kelley et al., 2020) agree that few classical parameters from an automatic weather station as inputs to a simple ANN structure (Kumar et al., 2010) may reasonably estimate the evolution of surface turbulent heat fluxes. Finally, in their review concerning ANN-based modeling of evapotranspiration, Kumar et al. (2010) highlighted one main limitation of the ANN estimation method: it is location-specific. Therefore, each automatic weather station deployed over a certain land cover type needs its own ANN configuration. The same conclusion was highlighted by Kelley et al. (2020).

The objective of our study is to test a low-cost and long-term way to monitor the surface turbulent heat flux heterogeneity over a landscape using a basic automatic weather station network, one eddy-covariance

station and ANNs. The eddy-covariance station is maintained at one location for a fixed time interval and then moved from one land-cover type to another. The final goal is that the meteorological conditions measured by the automatic weather stations over a surface cover feed a trained ANNs for this surface to estimate its exchange with the atmosphere. In this study, two one-year datasets are used to simulate the deployment of this method before its real deployment during a one-year field campaign. These datasets include basic automatic weather station variables and surface turbulent heat fluxes (sensible heat flux and latent heat flux) measured over a meadow and a maize crop field at heights of 2.5 and 4.5 m, respectively. The meadow changes slowly during the year, whereas the crop shows rapid growth or abrupt changes when harvested. The rotation frequency of the eddy-covariance station from one instrumented surface to another is an important experimental setting because it determines the dataset used to train the ANNs. Several rotation frequencies will be tested on the two studied vegetation covers.

The conceptual experimental set-up and the datasets used to test the method are presented in Section 2. Then in Section 3, we evaluate the efficiency of the ANNs in estimating surface-atmosphere interactions according to the rotation period of the eddy-covariance station and the type of vegetated surfaces. ANNs performances are also analyzed according to the temporal resolution of the flux estimates. The discussion and conclusions are presented in Section 4.

## 2. Methodology and Data

### 2.1. Conceptual experimental set-up

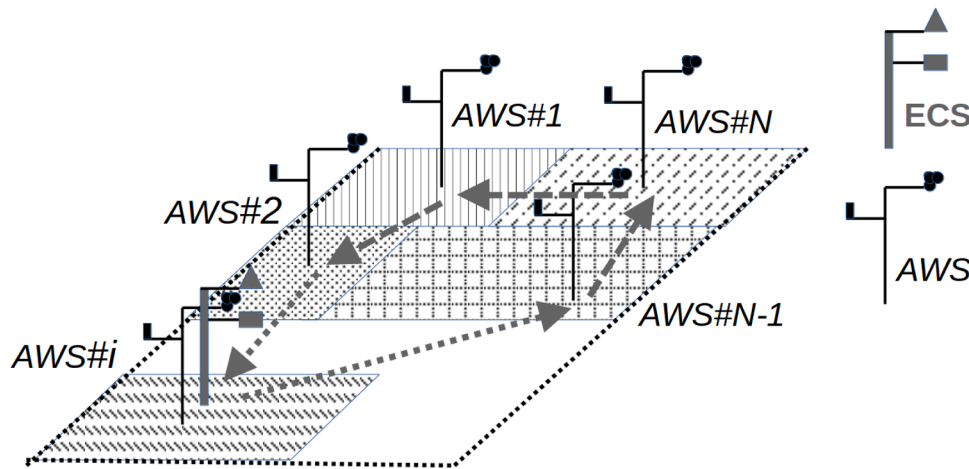
Given a heterogeneous landscape with  $N$  main surface covers (Fig. 1), the experimental set up would be composed of  $N$  Automatic Weather Stations (AWSs) distributed over the surfaces and one single Eddy-Covariance Station (ECS) moving from one AWS to the other with a rotation period to be determined. This combination of fixed AWSs and moving ECS should be implemented for at least one year to sample various meteorological and soil conditions over each of the surfaces. Because the ANN is specific to the instrumented surface, there will be  $N$  ANNs. The repeated time period during which the ECS stays close to each AWS is called the training-period ( $P_{training}$ ) for the ANN. During that period, the meteorological variables from the AWS are used as input variables, and, the surface fluxes from the ECS are used as control variables for the ANN training. The rotation-period ( $P_{rotation}$ ) is defined as the time for the ECS to visit all the AWSs so that it is the sum over  $N$   $P_{training}$  (Eq. 1). As an example, a  $P_{training}$  of one-week would lead to a  $P_{rotation}$  of  $N$  weeks. Successive rotations throughout one year would be applied so that as many vegetation stages and meteorological conditions as possible are sampled. The entire year is then decomposed in repeated  $P_{rotation}$  at the period of  $P_{rotation}$ .

$$P_{rotation} = \sum_{i=1}^N (P_{training})_{AWS\#i} \quad (1)$$

One of the objectives of this study is to examine how long the  $P_{training}$  and the  $P_{rotation}$  should be to correctly sample the vegetation and meteorological conditions so that the ANN of each surface type (slowly evolving for the meadow, and highly changing for the crop field) is correctly trained.

### 2.2. The Artificial Neural Network

ANNs, consisting of interconnected mathematical neurons as developed by Rosenblatt (1960), are widely used in data-driven modelling as they frequently outperform other approaches, especially in the parametrization of surface fluxes (Leufen and Schädler, 2019). For our purposes, we use the most basic type of ANN, the multi-layer perceptron also known as feed-forward network, in which neurons are arranged in fully interconnected sequential layers; one input layer to receive the



**Fig. 1.** Conceptual experimental set-up of  $N$  automatic weather stations (AWS) distributed on the  $N$  main land covers composing a landscape. The Eddy-Covariance Station (ECS) will be regularly moved from one AWS to the other so that concurrent dataset of AWS and ECS are available for ANNs training for each land cover.

predictors, one output layer to deliver the response and at least one intermediate layer in between, the so-called hidden layer. As a supervised ANN, it is trained using paired inputs (here the meteorological parameters) and targets (here the surface latent and sensible heat fluxes), and its internal parameters (weights and biases) are adjusted accordingly. The neuron output is calculated by adding a bias to the weighted sum and then applying an activation function to the result.

There is no well-established rule that defines the hyperparameters of ANNs, i.e., the number of hidden layers, their number of neurons, and their activation function. Several sensitivity experiments are then necessary (Leufen and Schädler, 2019). In this study, all the ANNs were implemented using tensorflow package (2.18.1 version). Following the previous studies of Leufen and Schädler (2019) and Zouzoua et al. (2024), the hyperbolic tangent and linear functions are used as activation functions for the hidden and output layers, respectively. The neurons in the output layers are fixed to two, one for each flux. Eventually, an adequate topography for the hidden layer was found through sensitivity experiments carried out with Keras auto-tuner (1.4.7 version; O'Malley et al. (2019)). It consists of three hidden layers, each with 10 neurons. Before training the ANNs, the input and target data are normalized to fit a centered-reduced law.

The ANN's performance is quantified in terms of Pearson's correlation coefficient ( $r$ ), root-mean-squared-error ( $RMSE$ ) and slope of the linear regression between observed and ANN-based surface turbulent heat flux. In order to take into account the effect of random initialization of the network weights and biases on the ANN estimates of surface flux, all the results in this study are presented as a statistic (median and standard deviation) over 50 trained ANNs with the same vector of input parameters. A sensitivity study was conducted and showed that 30 to 35 initialized runs with the same ANN structure and different random weights and biases are necessary to stabilize the performance of the ANNs for each scenario considered here (c.f., Section 2.3). However, a few runs still showed mediocre results, with scores at the extremes of the distribution, i.e. beyond the mean plus three times the standard deviation. We decided to repeat 50 identical runs to ensure that the ANNs performances are not substantially affected by these poorer runs.

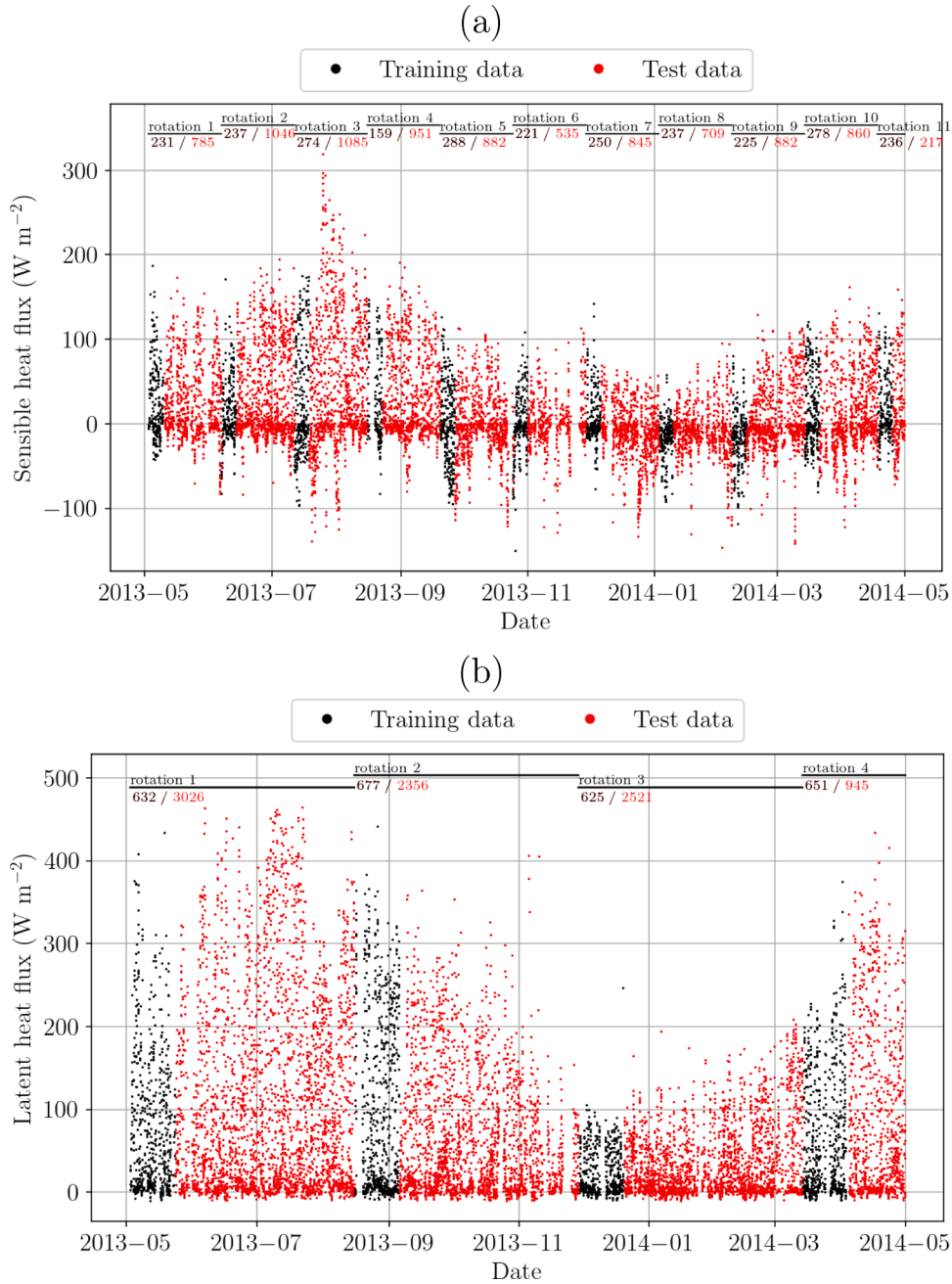
Our scientific objective is to evaluate the surface flux heterogeneity over a landscape. This means that we seek to evaluate the range of variation of the surface flux on time scales ranging from 10-days to monthly, rather than the exact surface flux at the temporal resolution at which it is calculated (30-minute sample). The choice of input variables is crucial for the ANNs to work properly. An accurate estimation of the fluxes every 30 minutes might be obtained using a large number of meteorological and satellite variables. Basically, the input must be related to the

targeted output. Accordingly, here we are tempted to use physical variables involved in the bulk formulation of surface turbulent heat fluxes. The input variables could therefore be (1) the vertical gradients of temperature and humidity, as well as the wind speed to define atmospheric air conditions, (2) the leaf area index for the vegetation state, (3) the soil temperature and moisture, and finally (4) the incoming solar radiation for the energy input to the system.

With such a set of inputs, the ANNs would mimic the physical relationship of the bulk formulae. However, in line with our technical objective, we try to make these estimations as cheap and simple as possible. The proposed method is thus based on the use of variables measured by basic meteorological stations, as those used in operational networks. Then, the ANNs link the basic meteorological characteristics of air near the surface to the turbulent heat fluxes, as a consequence and also a cause, rather than mimic the physical explanation of the flux as a bulk formulae would do.

Basic meteorological stations are usually equipped with sensors for the measurements of incoming shortwave radiation, horizontal wind components, air temperature, relative humidity, pressure, and rainfall close to the surface. However, the stations used to test this method are equipped with either a net radiometer or an incoming solar radiometer. Various studies have shown that radiative information such as net radiation (Zouzoua et al., 2024), incoming shortwave radiation (Ravindran et al., 2021) or sunshine hours (Sattari et al., 2021) is crucial for the estimation of surface fluxes using an ANN-based method. Dimitriadou and Nikolakopoulos (2022) used the three parameters (among others) as inputs to estimate evapotranspiration and studied the normalized importance of the input variables in the model performance. They found the following ranking among these three parameters, from the most important parameter to the least important: sunshine hours, incoming solar radiation and net radiation. Sunshine hours and incoming solar radiation have an almost similar percentage of normalized importance (around 40%) while the percentage for the net radiation is slightly lower (around 25%). The three radiative parameters are, however, less important than the mean air temperature, the vapour pressure deficit and the wind speed at 2 m from the surface. Based on this information, we assume that using net radiation or incoming solar radiation will not substantially influence our network performances and is in line with the idea of using basic and available meteorological station parameters as inputs.

This study will then evaluate the possibility of estimating the range of variation of the surface convective flux along a seasonal cycle with ANNs trained with net or incoming solar radiation, horizontal wind components, temperature and water vapour content. To these meteorological parameters, we add two temporal parameters that correspond to the



**Fig. 2.** (a) One year sensible heat flux measured by the eddy-covariance station in the meadow with the training and the test datasets in case of scenario 1. The duration of the 10.4 rotations of scenario 1 are indicated by black horizontal segments. The ratio of the number of 30-min samples in training and test datasets are indicated for each rotation. (b) Same as (a) but for latent heat flux in case of scenario 3. Thus, only 3.5 black horizontal segments indicate the rotations throughout the year in this panel. Period without data are those with instrumental breakdowns or erroneous data.

cyclic hour of the day:

$$DT_{cos} = \cos\left(2\pi \times \frac{DT}{24}\right) \quad (2)$$

$$DT_{sin} = \sin\left(2\pi \times \frac{DT}{24}\right) \quad (3)$$

with  $DT$  the Decimal Time so that  $\frac{DT}{24}$  varies from 0 at 00:00 to 0.98 at 23:30 every day. Half-hourly values of  $DT$  are defined because surface turbulent heat fluxes are estimated from 30-min sampling, as explained in Section 2.3.

Several tests on the input variables were conducted to check their relevance in estimating the surface fluxes. For example, we tested the impact of the use of wind speed instead of horizontal wind components, which adds information about the air mass origin. The performance of

the ANNs was not improved significantly. However, a more heterogeneous footprint might support incorporating air mass provenance into the ANNs inputs. Even if it might be unexpected from a physical point of view, the two diurnal cycle coordinates (Eqs. 2 and 3), also used by Zouzoua et al. (2024) and Kelley and Paradyjak (2019), improve the ANNs performance. In addition to these temporal coordinates, Zouzoua et al. (2024) use two other similar parameters defining the seasonal cycle. However, including these seasonal coordinates in our input variables deteriorates the ANNs performance. This might be due to an over-representation of time (4 input parameters) compared to only five meteorological variables and because the data used in this study only covers one annual cycle. This illustrates that no rigid rules regulate the ANN use, thereby tests are necessary to verify the relevance of the input variables according to the dataset, the objectives and the desired accuracy.

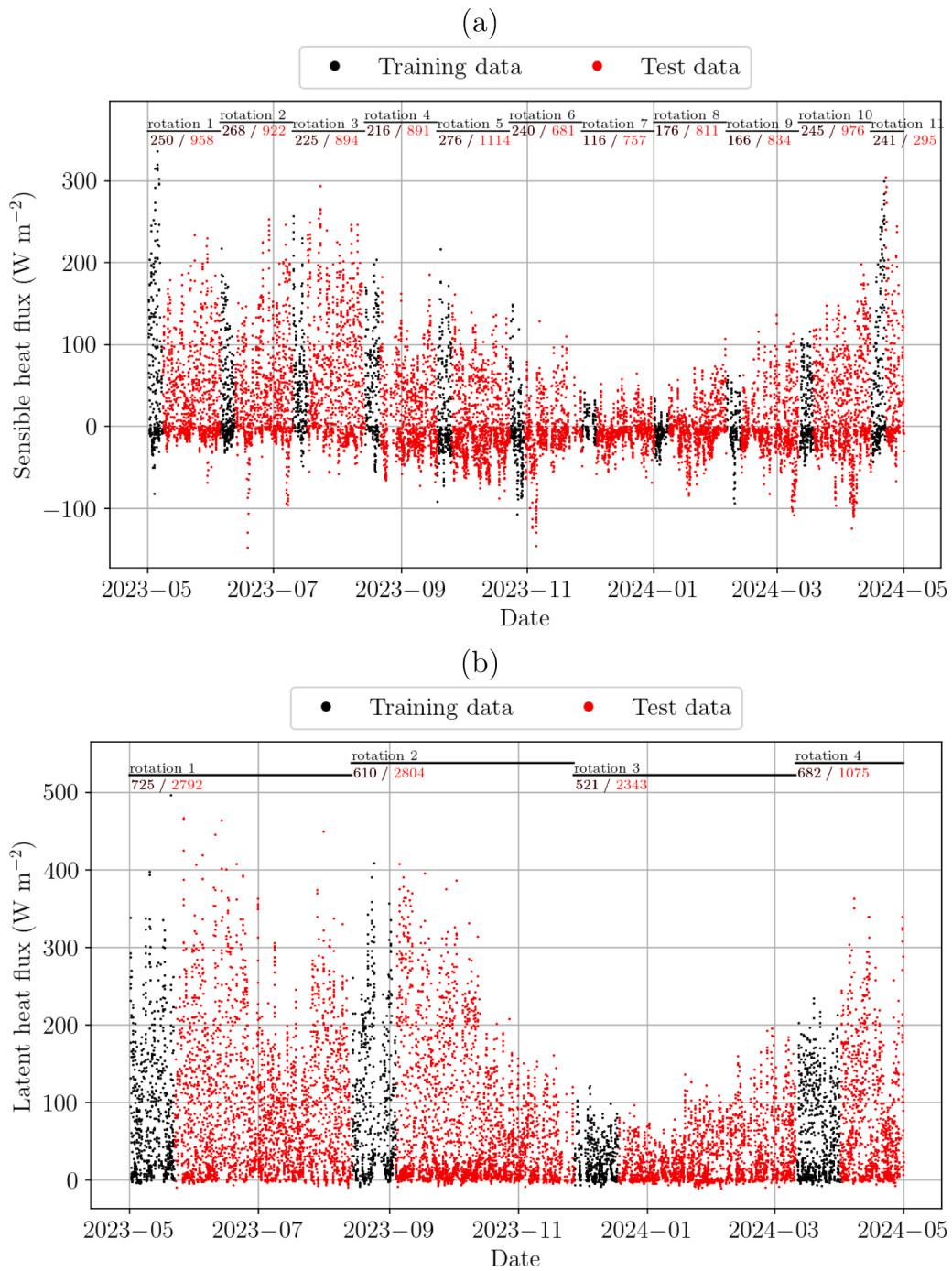


Fig. 3. Same as Figure 2 for the crop site.

### 2.3. ANN method evaluation and data set

To evaluate the method, we used meteorological variables (input variables) and turbulent surface turbulent heat fluxes (control variables) measured over one year in two fields: (1) a meadow that evolves slowly throughout the year, (2) a crop field in which the vegetation changes from grass during winter to maize during summer. A set of 50 ANNs are trained and tested for each of these two fields. Before we describe the measurement sites later in this section, the way these two one-year datasets are used to train and evaluate the ANNs is explained.

This full year of input and control variables measured at the same time in the two fields allows us to change the  $P_{training}$  and evaluate its im-

pact according to the vegetation cover. Each of the two one-year datasets is separated into a training and a test datasets. The training data are used for training (adjustment of the network weights) and validation (application of the trained network) of the ANN. The validation phase allows us to see how the network performs on new data, not used for the ANN training. In particular, during the validation, we search for potential network overfitting (i.e., the ANN has learned too “much” from its training and is not able to perform on new data). The data used for validation represent 10% of the training dataset. Once the ANN is trained and validated, it is applied to the test dataset to evaluate the ANN scores for data that are not used for training or validation. The way in which training and test datasets are built according to the rotation scenario is now discussed.

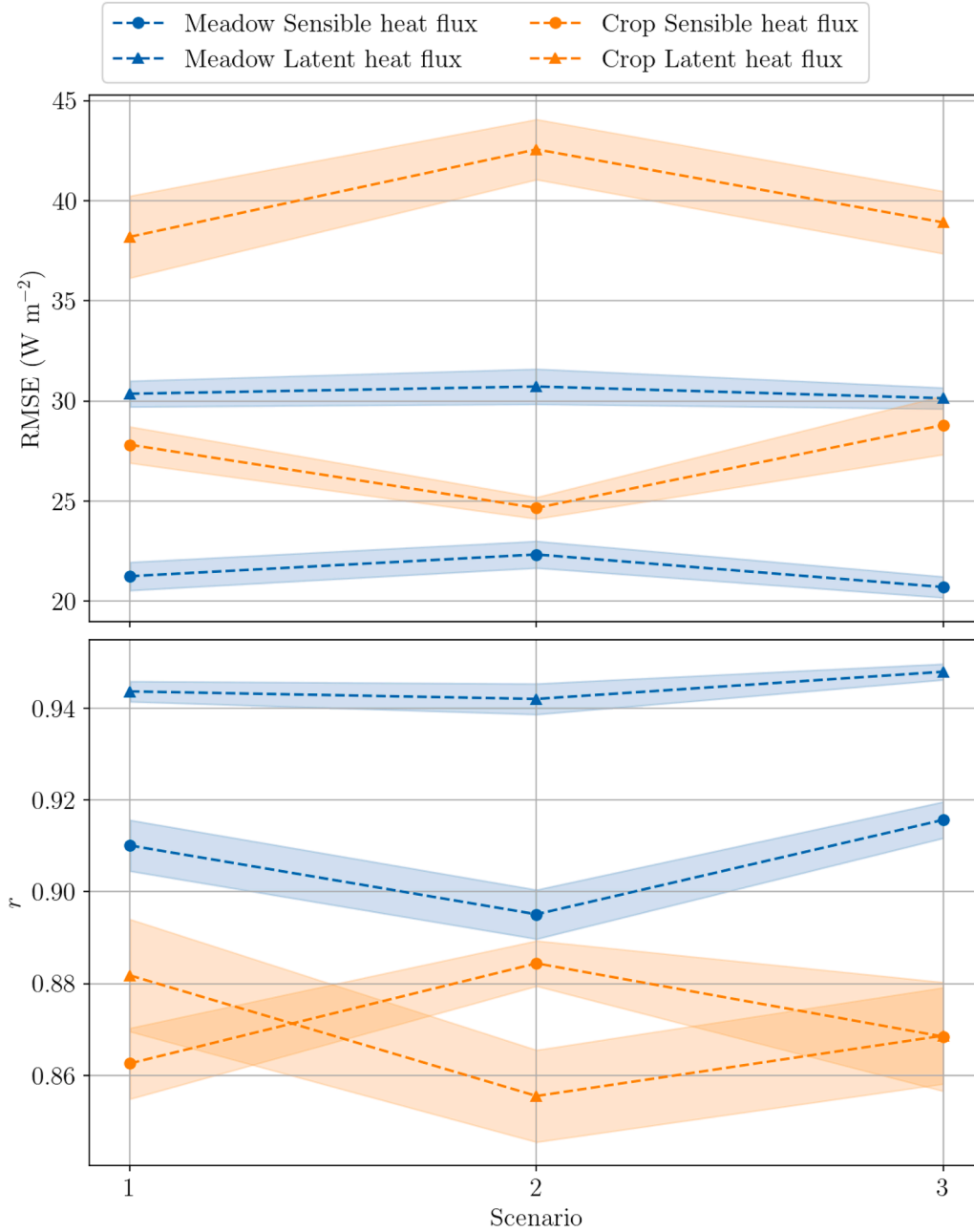


Fig. 4. Median and standard deviation of the scores  $RMSE$  and  $r$  of the 50 ANNs in simulating surface sensible and latent heat fluxes for the test dataset according to the three ECS rotation scenarios for the meadow and the crop field.

We arbitrarily assume that the landscape is characterized by five main surface covers ( $N = 5$ ), the meadow and the crop field being two of those. Three  $P_{training}$  are tested: 1 week, 2 weeks, and 3 weeks. The time in between two consecutive  $P_{training}$  in a given surface cover, is called the test-period ( $P_{test}$ ) because the dataset acquired during that period can be used to evaluate the ANN flux estimates.  $P_{rotation}$  is the sum of  $P_{training}$  and  $P_{test}$  and is repeated throughout the year. The three scenarios tested in this study can be summarized as follows:

- scenario-1:  $P_{training}$  of one week,  $P_{test}$  of 4 weeks, repeated 10.4 times over one year,
- scenario-2:  $P_{training}$  of two weeks,  $P_{test}$  of 8 weeks, repeated 5.2 times over one year,
- scenario-3:  $P_{training}$  of three weeks,  $P_{test}$  of 12 weeks, repeated 3.5 times over one year.

It should be noted that the year is not divisible by  $P_{rotation}$ , which explains the repetition of 10.4, 5.2 or 3.5 for the three scenarios. The minimum value of the  $P_{training}$  is one week. It is relevant to sample with high frequency the evolution of the vegetation state and the range of meteorological conditions throughout the year. However, moving the ECS is time consuming (at least one day of work), so that achieving this once a week throughout the year can be difficult to realize. In addition to the manpower issue, this scenario implies that there is no ECS measurements for one day per week, inducing a loss of 14% of the ECS data. The maximum value of three weeks for  $P_{training}$  is motivated by the fact that the seasonal change of vegetation and meteorological conditions must still be sampled frequently enough to catch the different stage of their evolution, especially for the crop field.  $P_{training}$  of one to three-week thus correspond to satisfying compromise.

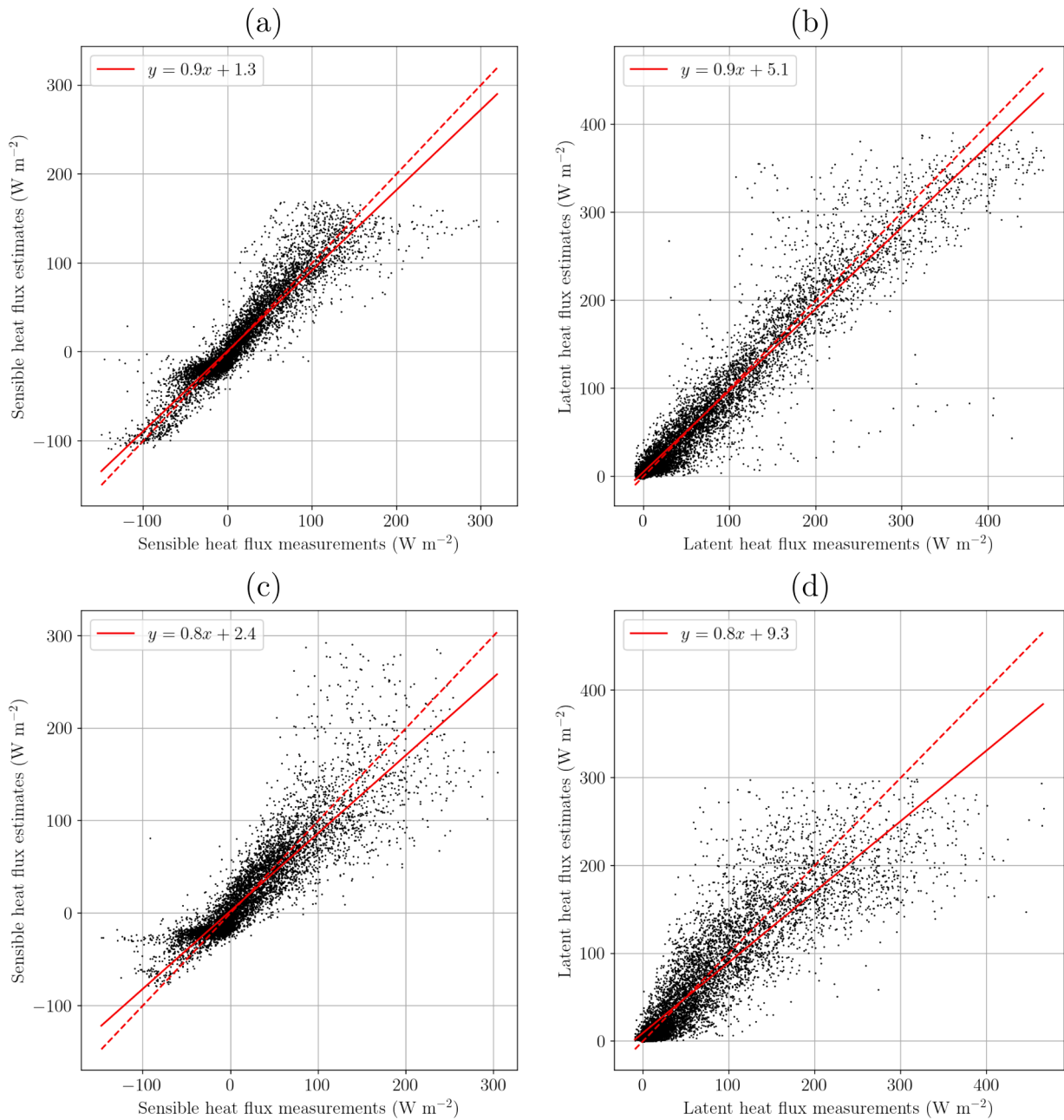


Fig. 5. Sensible heat (a) and latent heat fluxes (b) median estimate (of the 50 ANNs estimates) according to the fluxes calculated with eddy-covariance measurements for the meadow. (c) and (d) same as (a) and (b), respectively, for crop field. The 30-min samples of the test dataset of *S3.meadow* and *S3.crop* are used in this figure. The linear regression (continuous red line) and its equation are displayed in each panel and can be compared to the 1:1 line (dashed red line).

The meadow and crop datasets used in this study were acquired at the P2OA instrumented site (<https://p2oa.aeris-data.fr/>, Lothon et al. (2024)). The P2OA site is located in South-West France, on the Lanmezean Plateau in the northern front range of the Pyrénées mountains. The most frequent synoptic meteorological situations are (1) the westerly and north-westerly Atlantic flow, (2) the southerly winds often associated with Foehn situations, and, (3) anticyclonic conditions with weak northeasterly flow. Under anticyclonic situation or low synoptic forcing, the P2OA site is impacted by mountain breezes from the north-east for the anabatic wind and the south-east for the katabatic wind. The annual rainfall is about 1200 mm. All the instrumentation and meteorological conditions at the P2OA site are described in Lothon et al. (2024).

The meadow site is located in a cow pasture that maintains a regular height of vegetation throughout the year between 20 and 50 cm. For the

crop site, the AWS and EC stations were set up in the middle of a large area of maize cultivation during summer and grass during winter. However, this large area consists of several fields that were not all planted with maize at the same time, resulting in different stages of vegetation around the station (Table 1). Compared to the meadow site, the crop site produces difficulties for our ANNs method. Not only does the vegetation cover change over the year, it is not always fully homogeneous in the footprint of the eddy-covariance station. These two extreme configurations will allow us to compare the scores of the meadow's ANNs and the maize's ANNs.

Table 2 gives, for the meadow and crop sites, the measurement heights, the beginning and end dates of the datasets, the instrumentation, the measured parameters and the acquisition frequency. For the meadow station, ECS measurements are used to compute both

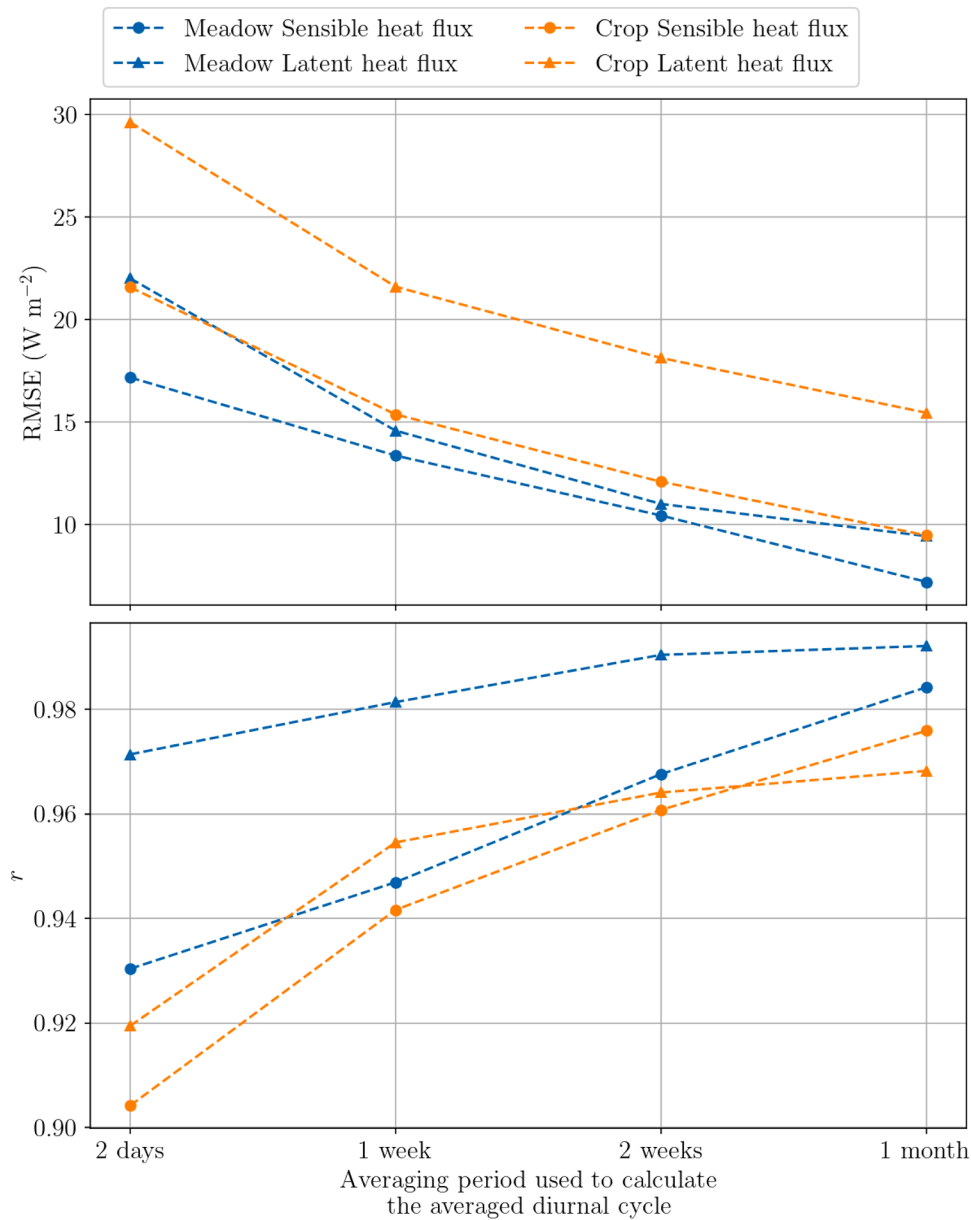


Fig. 6. ANN scores ( $RMSE$  and  $r$ ) at estimating average diurnal cycles of sensible and latent heat fluxes for meadow and crop field according to the averaging period used to calculate average diurnal cycles.

input variables and target/control variables which are indicated in the two last columns of Table 2. This was done because no AWS was associated with this EC station. On the contrary, for the crop field, an AWS was set up and provided the input parameters. Input and target/control parameters are then given by different instruments. From the measurements of the sonic anemometer and the CO<sub>2</sub>/H<sub>2</sub>O analyser, 30-minute surface sensible ( $H$ ) and latent ( $LE$ ) heat fluxes have been estimated using the eddy-covariance method EddyPro® Software (Version 6.2.0) from LI-COR (<https://www.licor.com/env/support/EddyPro/topics/whats-new.html>). Consistently with the time resolution of the surface flux, the input variables are 30-minute averaged.

Standard quality control procedures are applied. Rain events have been removed and flux calculations are flagged according to Mauder and Foken (2011). Only flux estimates with flags equal to 0 or 1 are used here, ensuring that basic hypotheses for flux calculation by eddy-covariance method are satisfied. Gap-filling procedures were not used to prevent introducing additional bias. In total, erroneous data and instru-

mentation break down affect 34.7% (34.2%) of the samples acquired on the meadow (crop), so that 11433 (11552) samples over 17519 (17567) remain after data cleaning and selection. There are several reasons for the large number of rejected samples: (1) if only one parameter (input or control) is not available at time  $t$ , this sample cannot be used either for training the ANN or for its evaluation; (2) the measurement site is located in a rainy area, which affects the measurement of humidity by the gas analyzer and therefore the estimation of  $LE$ .

Figures 2 and 3 present the temporal evolution of  $H$  and  $LE$  throughout the year in meadow and crop field, respectively. Training data and test data split is displayed in case of scenario 1 (one-week  $P_{training}$ ) in Figures 2a and 3a, and in case of scenario 3 (three-week  $P_{training}$ ) in Figures 2b and 3b. Training and test datasets do not rely on continuous 30 minute samples as would other methods because samples (in time) are treated independently from each other in the ANN method. These figures highlight how differently the training data are distributed throughout the year between the two scenarios. This has two important

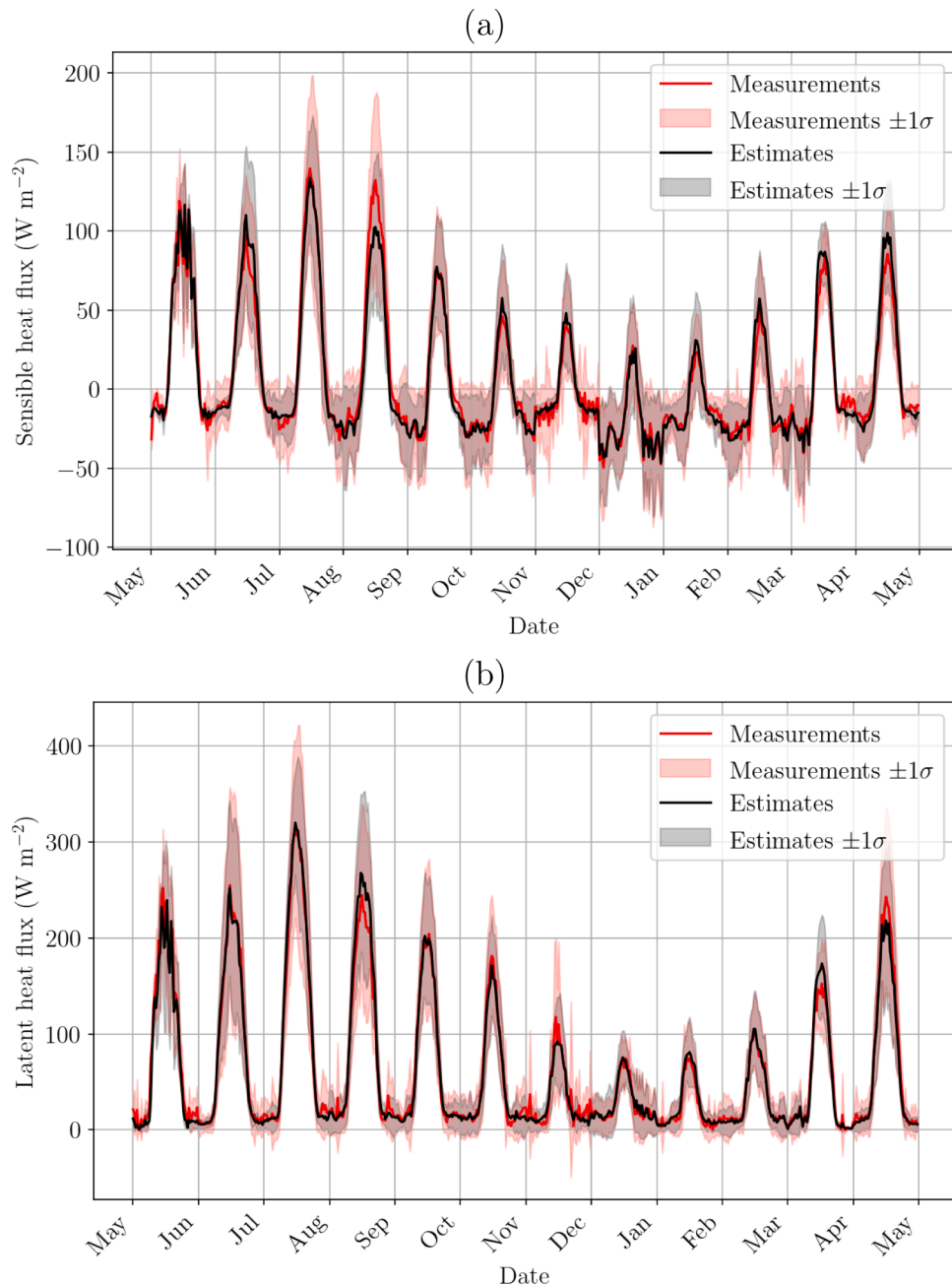


Fig. 7. Monthly averaged diurnal cycle and its standard deviation for the eddy-covariance measurements and the median ANN estimates of (a) sensible and (b) latent heat flux for *S3.meadow*.

consequences. First, the training dataset includes different conditions in terms of meteorology, surface conditions and vegetation state. For example, it can be noted for the meadow (Fig. 2) that neither scenario allows for the inclusion of the high values of  $H$ , that occur in late July and early August, in the training data set. The poor representation of the most unstable conditions in the training dataset could imply weaker performance of the ANN during the summer. With regard to  $LE$ , the training dataset for scenario 1 includes the high values measured during rotations 1 and 2, while the training dataset for scenario 3 does not. This could imply a better performance of the ANN in estimating high  $LE$  values in case of scenario 1 than in scenario 3. The second consequence is the different partition of the 30-minute samples between the training and the test datasets according to the scenario, as shown in Table 3. These differences should be considered when the ANN's performances are compared according to the scenario.

### 3. Results

The ANN scores for the three scenarios and the two vegetation covers are now analyzed. The  $r$  and  $RMSE$  between the measured and each of the 50 ANN estimates of the surface turbulent heat fluxes are calculated for the training and test datasets. The median values of the scores are presented in Table 4 and displayed, with their standard deviations, in Figure 4. The name of the test is composed of the scenario number and the vegetation cover on which this test is performed. As an example, *S1.meadow* is the test for one-week training-period applied on the meadow dataset.

Table 4 and Figure 4 show relatively good median scores and small standard deviations, with  $r$  above 0.86 and  $RMSE$  below  $42 \text{ W m}^{-2}$ , irrespective of the scenario and the vegetation cover. These results are reassuring because, at first sight, they suggest a fairly robust method.

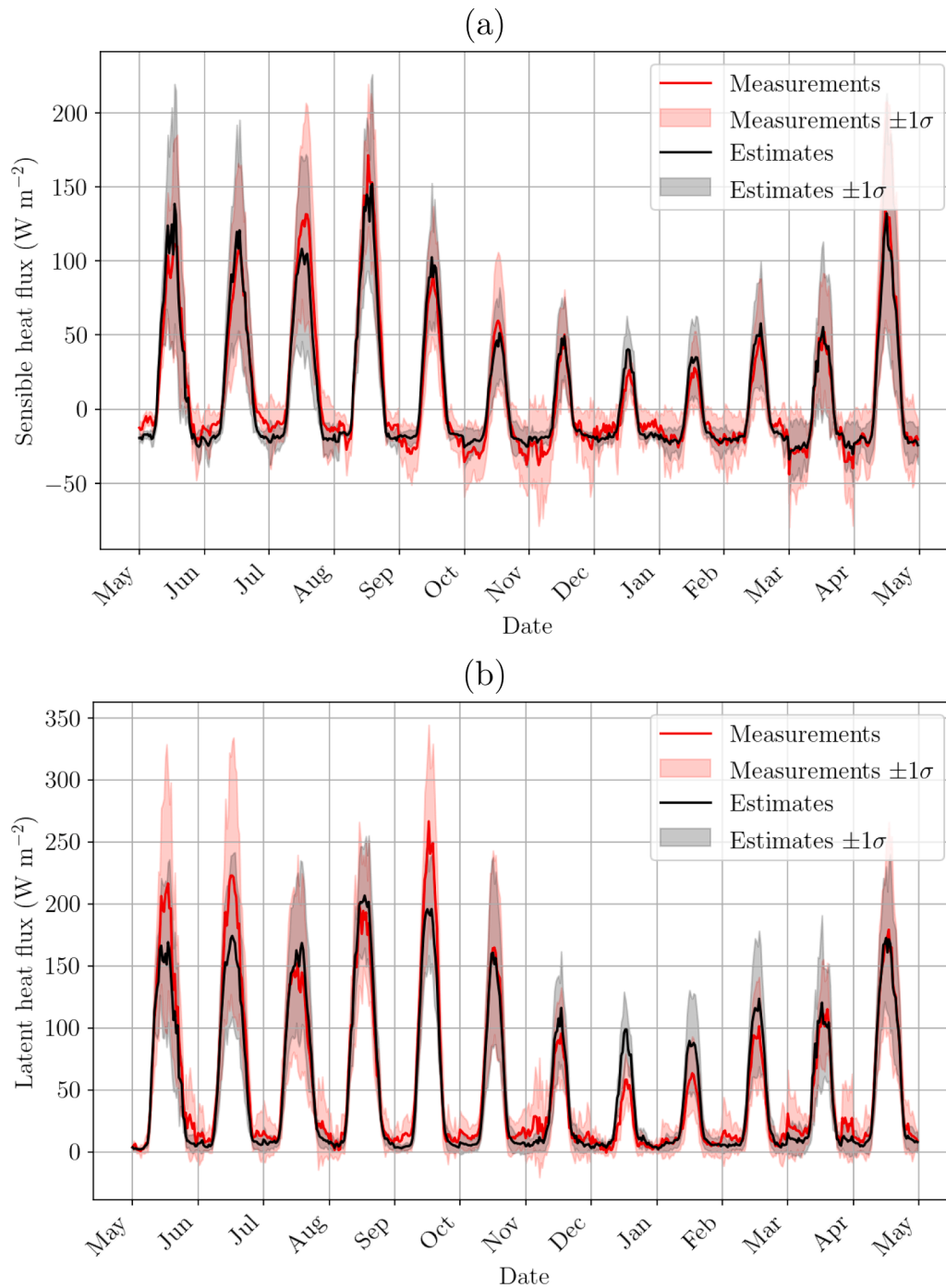


Fig. 8. Same as Figure 7 for *S3.crop*.

However, there is more to say about the variability of the scores depending on the scenario and vegetation cover.

### 3.1. Training-period effect

The test on the ECS rotation scenario should allow us to decide which  $P_{\text{training}}$  is acceptable in terms of ANN scores and manpower involvement for ECS moving from one AWS to the other. As shown in Table 4 and Figure 4, the scores for the test dataset do not change drastically from one ECS rotation scenario to another. This is particularly the case for the meadow site, while opposite weak score trends are observed for  $H$  and  $LE$  in the crop field. Whereas *S2.crop* has the largest  $RMSE$  for  $LE$  compared to *S1.crop* and *S3.crop*, it is the opposite for  $H$ . It is im-

portant to note that the scores differences are inconsequential and do not depend on the length of the training dataset (Table 3). Scenario 2 has the training dataset with the largest number of 30-min samples but does not systematically have the best scores for  $H$  and  $LE$  estimates. To understand the impact of sampled conditions (not shown) in the training dataset, the  $P_{\text{training}}$  was delayed a few days to include, in the training dataset, some of the highest values of one of the surface turbulent heat flux. The performance of the ANNs are then slightly improved for the flux considered once applied to the test dataset. However, this result contains several biases. The improvement of the ANNs performance concerns only one flux.  $H$  and  $LE$  being linked by the surface energy balance equation, favoring a good representation of high  $H$  values in the training dataset reduces the representation of high  $LE$  values.

**Table 1**

Vegetation cover and height in the fields surrounding the measurement station at crop site throughout the year.

Date	Vegetation		Comment
	cover	height (cm)	
2024/06/08	grass maize	5 < 10	first field planted in maize
2024/08/03	maize	10 to 200	according to fields
2024/09/12	maize	250	in all fields
2024/09/27	harvested field maize	0 250	first field harvested
2024/11/29	bare soil	0	all fields harvested
2025/02/14	grass	5 to 10	in all fields

Furthermore, as the dataset is limited to one year in this method, if high values are placed in the training dataset, they will be less present in the test dataset and therefore the ANN's ability to estimate them will be less evaluated.

As mentioned above, one-week  $P_{training}$  is practically very time consuming and one day of measurements is lost every seven days. Thus, two or three-weeks  $P_{training}$  seem to be a good compromise for both meadow and crop. One could have expected that a shorter  $P_{training}$  would be able to catch rapid growth or drastic vegetation changes at the crop site and would have improved the scores. This is clearly not the case. We will now analyze the ANN scores according to the vegetation cover based only on scenario 3.

### 3.2. Vegetation effect

The comparison between the two vegetation covers shows a loss of ANN performance for the crop compared to the meadow (Table 4 and Figure 4). For the test dataset of scenario 3,  $r$  is reduced from values above 0.92 for meadow to 0.87 for crop, and  $RMSE$  increases from 20.7 above meadow to 28.8  $W m^{-2}$  above crop for  $H$  and from 30.1 to 38.9  $W m^{-2}$  for  $LE$ .

Figure 5 illustrates a comparison between the median of the 50 ANN estimates and the measured heat flux for the  $S3.meadow$  and  $S3.crop$  test datasets. This figure highlights the ability of the ANN to simulate the different ranges of flux values. The sensible heat flux in  $S3.meadow$  (Fig. 5a) is underestimated by the ANN in the case of unstable conditions (values above  $150 W m^{-2}$ ). The overestimation of  $LE$  for such conditions is also observed but not systematically (Fig. 5b). This different ability of the ANN to represent the two heat flux for unstable conditions is probably related to the low occurrence of high values of  $H$  and the quite high occurrence of high values of  $LE$  in the training dataset in June and July (Fig. 2). Some very important underestimations of  $LE$  by the ANNs are also possible. Different ANN behaviors can be noticed for  $S3.crop$  compared to  $S3.meadow$ . In addition to lower scores as indicated in Table 4, the ANN estimates are 20% underestimated on average, with slopes of the linear regression for  $H$  and  $LE$  equal to 0.8 compared to 0.9 for the meadow. The ANN does not underestimate the high values of  $H$ , but the results are scattered above  $100 W m^{-2}$ , and  $LE$  is often underestimated for values above  $200 W m^{-2}$ . These behaviors could be partly explained by the highest values of  $H$  (early May) included in the training dataset (rotation 1), while the highest values of  $LE$  (June) are not (Fig. 3).

ANN seems to make the average of surface-atmosphere interactions of different vegetation covers that follow each other in the field throughout the year (Table 1). The ANN estimates of the flux are then what an average vegetation between bare soil, grass and maize would give. High values of  $LE$  cannot be accurately estimated because the average vegetation cannot transpire as much as fully grown maize.

### 3.3. Averaged diurnal cycle

As already mentioned, this method is set up to monitor the surface flux heterogeneity of a landscape. This means that at the scale of a model grid, which can vary from  $1km \times 1km$  to  $50km \times 50km$  for weather forecast and climate numerical models respectively, the range of variation of the surface flux must be determined. Therefore, a sufficiently accurate mean diurnal cycle with its standard deviation is required. The period over which all daily diurnal cycles are averaged to compute the mean diurnal cycle is called the averaging period. The choice of its length is the result of a balance between the accuracy and the temporal frequency

**Table 2**

Sensors, measured variables and acquisition frequency for the instrumented towers in the meadow and the crop field. Height measurements of the sensors and start and end dates of the two datasets are given. Input and target/control variables used for the ANN method are indicated in the two last columns. Both input and target/control variables have a 30-minute temporal resolution.

Sensor	Measured var.	Acq. freq. (Hz)	Input var.	Target-Control var.
<i>Meadow station - 2.5 m height - 1 May 2013 to 30 April 2014</i>				
NR lite (Kipp Zonen)	Net radiation $W m^{-2}$	0.2	Net radiation	
Sonic anemometer (CSAT Campbell Sci.)	Air temperature ( $^{\circ}C$ )	20	Temperature	Sensible and Latent heat flux ( $H, LE$ )
	3 wind components ( $m s^{-1}$ )	20	Hor. wind comp.	
CO <sub>2</sub> /H <sub>2</sub> O analyser (7500A Licor)	Air humidity ( $mmole m^{-3}$ )	20	Air humidity	( $W m^{-2}$ )
<i>Crop station - 4.5 m height - 1 May 2023 to 30 April 2024</i>				
GHIS320 (Campbell Sci.)	Incoming solar rad. ( $W m^{-2}$ )	0.2	Incoming solar rad.	
2D Sonic anemometer (windsonic4 Campbell Sci.)	Hor. wind comp. ( $m s^{-1}$ )	0.2	Hor. wind comp.	
Thermometer-hygrometer (windsonic4 Campbell Sci.)	temp. and rel. hum. ( $^{\circ}C, \%$ )	0.2	Air temp. and humidity	
Sonic anemometer and hygrometer (Irgason, Campbell Sci.)	3D wind comp. Air temp. and hum. ( $m s^{-1}, ^{\circ}C, mmole m^{-3}$ )	10 10		$H$ and $LE$ ( $W m^{-2}$ )

**Table 3**

Number of 30-minute samples in the training and test datasets for the three scenarios and for the two vegetation covers. Scenario 1, 2 and 3 are for  $P_{training}$  of one week, two and three weeks, respectively.

	Training		Test	
	meadow / crop	meadow / crop	meadow / crop	meadow / crop
Scenario 1	2528 / 2557	8905 / 9406		
Scenario 2	2906 / 2627	8527 / 9336		
Scenario 3	2597 / 2439	8836 / 9524		

**Table 4**

Median scores of the 50 ANNs estimates of the surface turbulent heat flux, for train and test datasets for the meadow and the crop field. The letter  $S$  composing the name of the tests stands for *Scenario* and the number corresponds to the scenario number.

	Sensible heat flux		Latent heat flux	
	$RMSE (W m^{-2})$	$r$	$RMSE (W m^{-2})$	$r$
	Train/Test	Train/Test	Train/Test	Train/Test
S1.meadow	10.7 / 21.5	0.97 / 0.91	18.6 / 30.4	0.98 / 0.94
S1.crop	17.6 / 27.9	0.96 / 0.86	23.4 / 38.4	0.94 / 0.88
S2.meadow	11.0 / 22.3	0.98 / 0.89	18.9 / 30.8	0.98 / 0.94
S2.crop	18.6 / 24.7	0.96 / 0.88	22.7 / 42.6	0.94 / 0.86
S3.meadow	11.2 / 20.7	0.97 / 0.92	19.8 / 30.1	0.97 / 0.95
S3.crop	16.0 / 28.8	0.96 / 0.87	23.1 / 38.9	0.95 / 0.87

of the flux estimates by the ANN. The accuracy (temporal frequency) increases (decreases) with the lengthening of the averaging period.

The evolution of ANN scores for *S3.meadow* and *S3.crop* according to the averaging period used to compute the mean diurnal cycle is presented in Figure 6. Two days, 1 week, two weeks and 1 month are used to estimate the mean diurnal cycles for both the flux measurements and the ANN estimates before computing the  $RMSE$  and  $r$ . Figures 7 and 8 illustrate the mean diurnal cycles with an averaging period of one month. As expected, the scores are better than those presented in Table 4 for which 30-min samples are used without average. The longer the averaging period, the better the scores for the two vegetation covers. Figure 6 also suggests that the averaging period could be different depending on the vegetation cover and its complexity.

The seasonal evolution of the two fluxes on the meadow (Fig. 7a, b) is well represented by the ANN estimates ( $RMSE$  below  $10 W m^{-2}$  and  $r$  above 0.98, Fig 6). The diurnal cycle is also very well represented, even at night when complex variations of  $H$  occur. However, one can note bad estimations of the surface turbulent heat fluxes at midday in August, with an underestimation of  $H$  and an overestimation of  $LE$ . These discrepancies between measurements and ANN estimates median value are the direct consequences of the already discussed imperfect sampling of unstable conditions in the case of scenario 3 (Fig. 2), so that the ANN cannot correctly estimate the flux in such conditions.

The seasonal evolution of the crop-field fluxes (Fig. 8) is well represented by the ANN for  $H$  but much less for  $LE$ . The latter is underestimated by the ANN in May, June and September and overestimated from November to February, implying a  $RMSE$  larger than  $15 W m^{-2}$  (Fig. 6). The ANN, trained by a range of different vegetation states and types, may simulate an average vegetation. This leads to  $LE$  underestimation by the ANN when the maize in the field is growing and transpires more than the average vegetation, and  $LE$  overestimation by the ANN when the maize is harvested and transpires less than the average vegetation.

#### 4. Discussion and conclusion

Surface-atmosphere exchanges are key parameters in assessing and validating weather and climatic forecast models (Boone et al., 2009; Lohou et al., 2014; Couvreur et al., 2016). Documenting and under-

standing their evolution is crucial for improving weather and climate forecasting, especially when considering heterogeneous conditions in landscape (Bou-Zeid et al., 2020; Shen et al., 2016) and meteorological conditions. To efficiently estimate surface-atmosphere exchanges in a considered area, a network of measurement stations is required. Eddy-covariance is a widely used method to locally measure surface-atmosphere exchanges. However, this method is based on fast response sensors which are rather expensive.

In this study, we applied a new method based on neural networks to estimate sensible and latent heat fluxes at lower cost, using simple automatic weather station parameters. The goal was to define a deployment plan to perform long-term monitoring of the surface turbulent heat fluxes heterogeneity over a landscape with an Automatic Weather Station (AWS) network and trained artificial neural networks (ANN). One eddy-covariance station (ECS) is needed to train the ANNs, but once the ANNs are well trained, the ANNs estimated surface turbulent heat fluxes rely only on the network of AWSs. The ANNs parameterizations depend on the land cover instrumented: one ANN for one AWS (Kumar et al., 2010; Kelley and Pardyjak, 2019). This is interesting from a financial point of view because only one ECS is needed to train several ANNs. In this study (1) we simulated the rotation of the ECS as if we needed to train five AWSs and, (2) we investigated the efficiency of the ANNs at estimating the sensible and latent heat flux on a meadow and a crop field. The meadow changes slowly during the year, whereas the crop shows rapid growth or abrupt changes when harvested. The use of these very different vegetation covers allows us to test this method in quite easy, with the meadow, and challenging, with the crop, conditions.

The one-year datasets over the meadow and the crop field were split in training and test datasets according to the frequency at which the ECS is moved from one AWS to another. This determines the sampling of the data included in the training dataset and then the ANN's ability to estimate the flux in all conditions (meteorological and vegetation state). The more representative the training dataset is of the variability of conditions, the better the ANN scores (Kelley and Pardyjak, 2019; Kelley, 2020). Three scenario were tested in this study: one-, two- and three-week rotations.

The results show that the use of AWS associated with an ANN allows to estimate sensible and latent heat flux with  $RMSE$  below  $42 W m^{-2}$  and Person correlation coefficient above 0.86 when 30-minute ANN estimates are compared to Eddy-covariance measurements, irrespective of the scenario and the vegetation cover. However, the objective of implementing this method is not to determine flux every 30 minutes, but rather to provide information on the variability of seasonal and intra-seasonal changes of flux across a heterogeneous landscape. This allows averages to be calculated that will significantly reduce the uncertainty of ANN estimates. The  $RMSE$  is divided by three when the diurnal cycle is monthly averaged. This averaging period can also be adapted to the type of vegetation and its evolution throughout the year.

Concerning the scenario, inconsequential differences were noticed in the ANN scores between the different scenarios. One could have expected that a more rapid rotation of the ECS on the different AWS would allow to include in the training dataset larger meteorological conditions range, rapid growth or drastic vegetation changes. However, ANN scores do not vary significantly for meadow and crop. Moving the station every one-week is logistically demanding when considering a one-year-long campaign. The two- or three-week rotation periods then seem to be a good compromise between good scores and logistically manageable deployment conditions.

The comparison between the two vegetation covers shows a loss of ANN performance for the crop compared to the meadow. The training set for the meadow consists only of data resulting from surface-atmosphere interactions on a single type of vegetation cover, which changes little throughout the year. The meadow training dataset only needs to sample as many weather conditions as possible. The ANN scores are therefore very good. Regarding the crop site, maize is planted in

spring, harvested in autumn, and grass is sown in winter. Between these two periods, the field goes through a stage of bare soil. The ECS rotation is carried out independently of these different cultivation stages, and the training dataset mixes different crops and different weather conditions. Therefore, it was expected that the ANN would perform less well than for the meadow, with an increase of the *RMSE* up to 40 % and a decrease of the Pearson correlation coefficient up to 8 % for 30-minute ANN estimates.

In some way, various improvements could be applied to get more robust results for crop field. For example, adding humidity or temperature gradients, or introducing larger training periods by buying more ECS to better sample the yearly evolution of the surface fluxes. However, we would like to remind that the main objective here was to monitor the evolution of several types of land cover at the lowest possible cost. We are not expecting perfect correlation between 30-min surface fluxes measurements and ANN estimated fluxes, as long as the overall yearly evolution is well represented. Therefore, adding complexity to the instrumentation deployed by adding more ECSs or sensors goes against this main objective.

However, two types of improvement could be considered. Firstly, adding information on the state of the vegetation cover might help overcome the issue of strongly evolving vegetation. To not increase the instrumentation cost with new sensors, remote sensing data could be used. For example, the Leaf Area Index is a commonly used remote sensing product that can be added as an input parameter to follow the vegetation evolution. This could help the ANN to distinguish between the different types of vegetation that grow in succession in a field. Remotely sensed data products would require considering scale-dependent effects when translating intermittent, gridded data to continuous, ground-based observations. Secondly, a transfer learning method could be applied. It consists of two steps. First, the ANN is trained using a large database acquired on the same type of agricultural field as the one studied. Next, the ANN is retrained, so that the weights and biases of the last hidden and output layers adapt to the shorter dataset being studied. Some large data bases for different crops field are available on ICOS ERIC database for example (<https://www.icos-cp.eu/>). Attention should be paid to weather conditions, which should not differ too greatly between the datasets used for the first and second steps.

Based on these results, the method was tested during the third and last Enhance Observation Period (EOP) of the Model and Observation for Surface-Atmosphere Interactions over heterogeneous landscape project (MOSAI project, (Lohou et al., 2025)). Three AWSs were deployed during this field campaign, along with one ECS moving from one station to another every three weeks. The three land covers instrumented were: meadow, summer (maize) and winter (wheat) crop fields. This experimental campaign will help us investigate the two major issues concerning this method mentioned above: (1) using remote sensing products to add an input parameter concerning the vegetation state, and (2) applying the transfer learning method.

#### CRedit authorship contribution statement

**Mathilde Jomé:** Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Conceptualization; **Fabienne Lohou:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Funding acquisition, Data curation, Conceptualization; **Marie Lothon:** Writing – review & editing, Methodology, Funding acquisition; **Lilane Fautrez:** Formal analysis; **Antoine Vial:** Resources, Data curation; **Solène Derrien:** Resources, Data curation; **Antoine Royer:** Visualization, Validation, Methodology, Data curation; **Maurin Zouzoua:** Writing – review & editing, Methodology, Investigation; **Eric Pardyjak:** Writing – review & editing, Software, Resources, Methodology, Conceptualization; **Jason Kelley:** Writing – review & editing, Software, Resources, Methodology, Conceptualization; **Sophie Bastin:** Writing – review & editing.

#### Data availability

Data will be made available on request.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

1. The experimental equipment (meteorological station network and eddy-covariance station) was purchased as part of the OBS4CLIM project which receives government funding managed by the French National Research Agency under France 2030, reference number 'ANR-21-ESRE-0013'.

2. The method proposed in this study was tested in the context of MOSAI project funded by French National Research Agency (Projet-ANR-20-CE01-0018) and took advantage of long-term measurements acquired at P2OA experimental site which is part of the ACTRIS (Aerosol, Clouds and Trace Gases Research Infrastructure; <https://www.actris.fr>) research infrastructure.

3. AERIS data center (<https://www.aeris-data.fr/>) is in charge of the data distribution and of the website of P2OA experimental site (<https://p2oa.aeris-data.fr/>) and MOSAI project (<https://mosai.aeris-data.fr>).

#### References

- van de Boer, A., Moene, A.F., Graf, A., Schüttemeyer, D., Simmer, C., 2014. Surface heterogeneity and its signature in higher-order scalar similarity relationships. *Bound.-Layer Meteorol.* 152, 19–44. <https://doi.org/10.1007/s10546-014-9920-8>
- Boone, A., de Rosnay, P., Balsamo, G., Beljaars, A., Chopin, F., Decharme, B., Delire, C., Ducharme, A., Gascoin, S., Grippa, M., Guichard, F., Gusev, Y., Harris, P., Jarlan, L., Kergoat, L., Mougou, E., Nasonova, O., Norgaard, A., Orgeval, T., Otlé, C., Pocard-Leclercq, I., Polcher, J., Sandholt, I., Saux-Picart, S., Taylor, C., Xue, Y., 2009. The AMMA land surface model intercomparison project (ALMIP). *Bull. Am. Meteorol. Soc.* 90 (12), 1865–1880. <https://doi.org/10.1175/2009BAMS2786.1>
- Bou-Zeid, E., Anderson, W., Katul, G., Mahrt, L., 2020. The persistent challenge of surface heterogeneity in boundary-layer meteorology: A review. *Bound.-Layer Meteorol.* 177. <https://doi.org/10.1007/s10546-020-00551-8>
- Couvreur, F., Bazile, E., Canut, G., Seity, Y., Lothon, M., Lohou, F., Guichard, F., Nilsson, E., 2016. Boundary-layer turbulent processes and mesoscale variability represented by numerical weather prediction models during the BLLAST campaign. *Atmos. Chem. Phys.* 16 (14), 8983–9002. <https://doi.org/10.5194/acp-16-8983-2016>
- Detto, M., Katul, G., Mancini, M., Montaldo, N., Albertson, J., 2008. Surface heterogeneity and its signature in higher-order scalar similarity relationships. *Agric. Forest Meteorol.* 148, 902–916.
- Dimitriadou, S., Nikolakopoulos, K.G., 2022. Artificial neural networks for the prediction of the reference evapotranspiration of the peloponnese peninsula, greece. *Water* 14 (13). <https://doi.org/10.3390/w14132027>
- Foken, T., Aubinet, M., Leuning, R., 2012. The eddy-covariance method. pp. 1–19. [https://doi.org/10.1007/978-94-007-2351-1\\_1](https://doi.org/10.1007/978-94-007-2351-1_1)
- Foken, T., Wimmer, F., Mauder, M., Thomas, C., Liebethal, C., 2006. Some aspects of the energy balance closure problem. *Atmos. Chem. Phys.* 6, 4395–4402. <https://doi.org/10.5194/acp-6-4395-2006>
- Kelley, J., 2020. Assessment and correction of solar radiation measurements with simple neural networks. *Atmosphere* 11 (11). <https://doi.org/10.3390/atmos11111160>
- Kelley, J., McCauley, D.M., Alexander, G.A., Gray, W.F., Siegfried, R., Oldroyd, H.J., 2020. Using machine learning to integrate on-farm sensors and agro-meteorology networks into site-specific decision support. *Transactions of the ASABE* <https://api.semanticscholar.org/CorpusID:226670248>.
- Kelley, J., Pardyjak, E., 2019. Using neural networks to estimate site-specific crop evapotranspiration with low-cost sensors. *Agronomy* 9, 108. <https://doi.org/10.3390/agronomy9020108>
- Kumar, M., Raghunwansi, N., Singh, R., 2010. Artificial neural networks approach in evapotranspiration modeling: A review. *Irrigation Science* 29, 11–25. <https://doi.org/10.1007/s00271-010-0230-8>
- Kumar, M., Raghunwansi, N., Singh, R., Wallender, W., Pruitt, W., 2002. Estimating evapotranspiration using artificial neural networks. *Journal of Irrigation and Drainage Engineering-ASCE - J IRRIG DRAIN ENG-ASCE* 128. [https://doi.org/10.1061/\(ASCE\)0733-9437\(2002\)128:4\(224\)](https://doi.org/10.1061/(ASCE)0733-9437(2002)128:4(224))
- Leufen, L.H., Schädlér, G., 2019. Calculating the turbulent fluxes in the atmospheric surface layer with neural networks. *Geoscientific Model Development* 12 (5), 2033–2047. <https://doi.org/10.5194/gmd-12-2033-2019>
- Lohou, F., Kergoat, L., Guichard, F., Boone, A., Cappelaere, B., Cohard, J.-M., Demarty, J., Galle, S., Manuela, G., Peugeot, C., Ramier, D., Taylor, C., Timouk, F., 2014. Surface

- response to rain events throughout the west african monsoon. *Atmos. Chem. Phys.* 14. <https://doi.org/10.5194/acp-14-3883-2014>
- Lohou, F., Lothon, M., Bastin, S., Brut, A., Canut, G., Cohard, J.-M., Cheruy, F., Couvreux, F., Dupont, S., Lafont, S., Roehrig, R., Román-Cascón, C., Yagüe, C., Bernard, E., Boone, A., Dupont, J.-C., Feofilov, A., Fernandes, R., Grulois, M., Jomé, M., Maison, A., Zouzoua, M., 2025. Model and observation for surface-atmosphere interactions over heterogeneous landscape: MOSAI project. *J.E.M.S.* 3, 100019. <https://doi.org/10.1016/j.jemets.2025.100019>
- Lohou, F., Saïd, F., Lothon, M., Durand, P., Serça, D., 2010. Surface heterogeneity and its signature in higher-order scalar similarity relationships. *Bound.-Layer Meteorol.* 136, 1–23. <https://doi.org/10.1007/s10546-010-9493-0>
- Lothon, M., Gheusi, F., Lohou, F., Pont, V., Soula, S., Derrien, S., Bezombes, Y., Leclerc, E., Athier, G., Vial, A., Philibert, A., Campistron, B., Saïd, F., Jambert, C., Sonke, J., Amestoy, J., Bargain, E., Bosser, P., Boulanger, D., Bret, G., Bodichon, R., Cabanas, L., Canut, G., Estrampes, J.-B., Gardrat, E., Gomez Kuri, Z., Gueffier, J., Guesdon, F., Lopez, M., Masson, O., Meslin, P.-Y., Meyerfeld, Y., Pascal, N., Pique, E., Ramonet, M., Starck, F., Vidal, R., 2024. The pyrenean platform for observation of the atmosphere: Site, long-term dataset and science. *Atmos. Meas. Tech* 17, 6265–6300. <https://doi.org/10.5194/amt-2024-10>
- Mauder, M., Foken, T., 2011. Documentation and instruction manual of the eddy covariance software package TK2. Arbeitsergebnisse, Universität Bayreuth, Abteilung Mikrometeorologie, ISSN 1614-8916 46. <https://doi.org/10.5194/bg-5-451-2008>
- Moene, A.F., Michels, B.L., HOLTSLAG, B.A.M., 2006. Scaling variances of scalars in a convective boundary layer under different entrainment regimes. *Bound.-Layer Meteorol.* 120, 257–274. <https://doi.org/10.1007/s10546-006-9053-9>
- Nakamura, R., Mahrt, L., 2001. Similarity theory for local and spatially averaged momentum fluxes. *Agricult. Forest. Meteorol.* 108, 265–279.
- O'Malley, T., Bursztejn, E., Long, J., Chollet, F., Jin, H., Invernizzi, L., et al., 2019. Keras-tuner. <https://github.com/keras-team/keras-tuner>.
- Ravindran, S.M., Bhaskaran, S. K.M., Ambat, S. K.N., 2021. A deep neural network architecture to model reference evapotranspiration using a single input meteorological parameter. *Environmental Processes* 8 (4), 1567–1599. <https://doi.org/10.1007/s40710-021-00543-x>
- Rosenblatt, F., 1960. Perceptron simulation experiments. *Proceedings of the IRE* 48 (3), 301–309. <https://doi.org/10.1109/JRPROC.1960.287598>
- Sattari, M., Apaydin, H., Shahab, S.B., Mosavi, A., Prasad, R., 2021. Comparative analysis of kernel-based versus ANN and deep learning methods in monthly reference evapotranspiration estimation. *Hydrology and Earth System Sciences* 25, 603–618. <https://doi.org/10.5194/hess-25-603-2021>
- Shen, L., Sun, J., Yuan, R., Liu, P., 2016. Characteristics of secondary circulations in the convective boundary layer over two-dimensional heterogeneous surfaces. *J. Meteorol. Res.* 30 (6), 2198–0934. <https://doi.org/10.1007/s13351-016-6016-z>
- Zouzoua, M., Bastin, S., Chiriaco, M., Lohou, F., Lothon, M., Jome, M., Mallet, C., Barthes, L., Canut, G., 2024. Using a data-driven statistical model to better evaluate surface turbulent heat fluxes in weather and climate numerical models: a demonstration study. *EGU sphere* 2024, 1–44. <https://doi.org/10.5194/egusphere-2024-568>