Partitioning net carbon dioxide fluxes into photosynthesis and respiration using neural networks

Gianluca Tramontana1,2 | Mirco Migliavacca3 | Martin Jung3 | Markus Reichstein3 | Trevor F. Keenan4,5 | Gustav Camps-Valls2 | Jerome Ogee6 | Jochem Verrelst2 | Dario Papale1,7

1DIBAF, Department for Innovation in Biological, Agro-food and Forestry Systems, University of Tuscia, Viterbo, Italy
2Image Processing Laboratory (IPL), Parc Científic Universitat de València, Universitat de València, Paterna, Spain
3Max Planck Institute for Biogeochemistry, Jena, Germany
4Department of Environmental Science, Policy and Management, UC Berkeley, Berkeley, CA, USA
5Earth and Environmental Sciences Area, Lawrence Berkeley National Lab, Berkeley, CA, USA
6INRAE, UMR 1391 ISPA, Villenave d'Ornon, France
7Euro-Mediterranean Center on Climate Change (CMCC), Viterbo, Italy

Correspondence
Gianluca Tramontana, Image Processing Laboratory (IPL), Universitat de València, Parc Científic Universitat de València, Calle Catedrático José Beltrán, 2, 46980 Paterna (València), Spain.
Email: gianluca.tramontana@uv.es; g.tramontana@unitus.it

Funding Information
European Commission, Grant/Award Number: 640176, 755617, 647423 and 703944; NASA, Grant/Award Number: NNH17AE86I; the EU’s Horizon 2020 Programme, Grant/Award Number: GA 820852; the EU’s Horizon 2020 Programme, Grant/Award Number: GA 703944

Abstract
The eddy covariance (EC) technique is used to measure the net ecosystem exchange (NEE) of CO2 between ecosystems and the atmosphere, offering a unique opportunity to study ecosystem responses to climate change. NEE is the difference between the total CO2 release due to all respiration processes (RECO), and the gross carbon uptake by photosynthesis (GPP). These two gross CO2 fluxes are derived from EC measurements by applying partitioning methods that rely on physiologically based functional relationships with a limited number of environmental drivers. However, the partitioning methods applied in the global FLUXNET network of EC observations do not account for the multiple co-acting factors that modulate GPP and RECO flux dynamics. To overcome this limitation, we developed a hybrid data-driven approach based on combined neural networks (NNC-part). NNC-part incorporates process knowledge by introducing a photosynthetic response based on the light-use efficiency (LUE) concept, and uses a comprehensive dataset of soil and micrometeorological variables as fluxes drivers. We applied the method to 36 sites from the FLUXNET2015 dataset and found a high consistency in the results with those derived from other standard partitioning methods for both GPP ($R^2 > .94$) and RECO ($R^2 > .8$). High consistency was also found for (a) the diurnal and seasonal patterns of fluxes and (b) the ecosystem functional responses. NNC-part performed more realistic than the traditional methods for predicting additional patterns of gross CO2 fluxes, such as: (a) the GPP response to VPD, (b) direct effects of air temperature on GPP dynamics, (c) hysteresis in the diel cycle of gross CO2 fluxes, (d) the sensitivity of LUE to the diffuse to direct radiation ratio, and (e) the post rain respiration pulse after a long dry period. In conclusion, NNC-part is a valid data-driven approach to provide GPP and RECO estimates and complementary to the existing partitioning methods.

Keywords
carbon dioxide fluxes partitioning, ecosystem respiration (RECO), eddy covariance, gross primary production (GPP), machine learning, net ecosystem exchange, neural network
1 | INTRODUCTION

The eddy covariance (EC) technique offers a unique opportunity for monitoring carbon and energy exchanges between land ecosystems and the atmosphere (Baldocchi, 2003) allowing near-continuous measurements integrated at the ecosystem scale. The number of study sites equipped with EC systems has increased over the years (Baldocchi, 2014; Chu, Baldocchi, John, Wolf, & Reichstein, 2017; Pastorello et al., 2017), and now we estimate that more than 1,400 globally distributed sites (Chu et al., 2017) are monitoring the most representative land ecosystems in different climate conditions. Most of the EC study sites are organized in regional networks such ICOS, AmeriFlux, and AsiaFlux, and contribute to the global FLUXNET network (Baldocchi, 2014).

The EC method allows for the measurement of the net ecosystem exchange (NEE) which is the difference between two larger flux components: gross primary production (GPP) and ecosystem respiration (RECO). GPP is the gross amount of carbon uptake by photosynthesis from vegetation while RECO is the total carbon efflux by the respiration processes of all organisms. Estimating GPP and RECO is a key step to better understand the underlying mechanisms constraining ecosystem function. Moreover, GPP and RECO estimates from EC are useful for modeling, supporting process-based model parameterization and validation, data assimilation, plant trait retrieval by model inversion (Dutta, Schimel, Sun, van der Tol, & Franken, 2019; Pacheco-Labrador et al., 2019), upscaling (Jung et al., 2020; Tramontana et al., 2016), as well as photosynthesis estimates based on remote sensing (e.g., Arnone et al., 2008; Verrelst et al., 2016; Zhang et al., 2016).

The EC technique does not directly measure GPP and RECO, and for this reason, numerical methods (termed partitioning methods) have been proposed for estimating GPP and RECO from NEE measurements (e.g., Desai et al., 2008; Keenan et al., 2019; Lasslop et al., 2010; Reichstein et al., 2005; Sulman, Tyler Roman, Scanlon, Wang, & Novick, 2016). The most widely used approaches are based on the use of NEE measurements for fitting simple physically-based nonlinear relationships to estimate GPP and RECO using few meteorological drivers. These functional relationships are in general either light response functions linking global incoming radiation and GPP (Gilmanov et al., 2003), or respiration functions based on temperature (Reichstein et al., 2005), or also combinations of the two approaches (Keenan et al., 2019; Lasslop et al., 2010). The simple implementation and the robustness of the results (Lasslop et al., 2010) have led to their adoption as standard processing tools in the FLUXNET community (Pastorello et al., 2020, accepted).

However, these partitioning methods rely on important assumptions about the flux dynamics and their relationship with environmental drivers. Importantly, the assumed functional relations used can be similar to those applied in ecosystem models that make use of these data for their validation, creating a risk of circularity. In addition, although the functional relationships used in standard partitioning approaches are known to be valid at the organ level (where they can be measured), their direct application at the ecosystem spatial scale should be evaluated carefully (Medlyn, 1998; Medlyn et al., 2017; Musavi et al., 2016). Furthermore, to guarantee their wide applicability, partitioning methods use a limited number of drivers (i.e., air temperature, vapor pressure deficit, and global radiation). However, fluxes dynamics are also potentially affected by many important environmental factors, such as soil moisture, soil temperature, or the ratio of diffuse to direct radiation, among others, that are generally not considered in the partitioning methods (Lasslop et al., 2012; Wohlfarth & Galvagno, 2017). These limitations are partially compensated by the use of short temporal moving windows for parameters estimation, which takes more slowly changing factors indirectly into consideration (such as phenology, water, or substrate availability), but does not handle fast ecosystem responses well, like for example respiration pulses following rain events (Williams, Hanan, Scholes, & Kutsch, 2009).

Some studies have attempted to solve the limitations of FLUXNET standard partitioning methods by developing more comprehensive approaches. For instance, Scanlon and Kustas (2010) coupled CO2 and H2O fluxes dynamics in the flux variance similarity approach for simultaneously partitioning carbon and water fluxes. The method, although interesting, requires canopy scale estimates of water-use efficiency (WUE), which introduces assumptions and uncertainty. More recently, the estimation of gross CO2 fluxes from NEE has been approached using the EC method in combination with parallel measurements of a trace gas such as carbonyl sulfide (COS; Commane et al., 2015) or 13C isotopes (Ogée, Peylin, et al., 2003; Oikawa et al., 2017; Whelan et al., 2018), limiting, for now, their application to a restricted number of study sites. These methods, however, are all subject to assumptions that could affect the resulting estimates of GPP and RECO. A possible alternative approach, based on existing measurements and not subject to the limitation of the standard methods, could be provided by machine learning methods.

Machine learning methods are generally used to model underlying complex relationships linking a set of predictors with one or more outputs. The core characteristics of machine learning methods are that predictors are not prescribed and the relationships between input (drivers) and output (fluxes) are inferred from the data. Several studies have reported the capability of machine learning to reproduce complex patterns in ecological studies, and in particular in relation to EC measurements (e.g., see Moffat, Beckstein, Churkina, Mund, & Heimann, 2010; Moffat et al., 2007; Papale & Valentini, 2003; Reichstein et al., 2019; Tramontana et al., 2016).

In this study, we develop and test an empirical machine learning-based approach for retrieving GPP and RECO from NEE. The main motivation was to evaluate if a purely empirical approach could provide estimates of the two components without any predefined
relationships and with the flexibility to use multiple sets of meteorological and biological forcings.

In this experiment, we used artificial neural networks (ANNs), which have already been tested as a partitioning tool in a cross-site intercomparison among several partitioning methods (Desai et al., 2008; Oikawa et al., 2017), without appreciable differences with respect to the other methods. In those studies, the ANNs were trained using only nighttime measurements and then applied to extrapolate RECO during daytime. Here, we revisited the ANN implementation by proposing a new scheme in which the ANN is constrained using the measured net CO₂ fluxes to directly estimate the two components GPP and RECO.

Our aim was to design a machine learning-based method flexible enough to: (a) use a large set of drivers, accounting for the complexity of the GPP and RECO responses, (b) ensure general applicability of the method to allow ANN-based partitioning in the context of FLUXNET thanks to the use of generally measured variables. In order to evaluate the results, we compared the GPP and RECO produced by the ANN with the estimates obtained by two partitioning methods used in FLUXNET, by analyzing: (a) the consistency of the estimates at yearly, seasonal, daily, and (half) hourly time steps; (b) the dynamics of the seasonal and diurnal cycles of the estimated gross CO₂ fluxes; (c) the functional relationships between micrometeorological inputs and fluxes in output as reproduced by the ANN; and (d) the effectiveness of the proposed method to predict additional patterns of gross CO₂ fluxes not (or less) accounted for by the standard partitioning methods.

2 | MATERIALS AND METHODS

2.1 | ANN algorithm

ANNs are nonlinear and nonparametric methods for regression and classification that artificially emulate the functioning of a biological brain (Haykin, 1999). The base unit of an ANN is the "neuron" where the numeric information in the input is weighted, condensed, and transformed (by a linear/nonlinear transfer/activation function) to be transferred to other neurons. Neurons are organized in layers that are interconnected: the outputs of m neurons in one layer are the inputs for n neurons of the next layer and the signals are transferred through connections associated with multiplicative weights. The learning (or training) process of an ANN consists of adjusting the weights of the network on the basis of specific examples provided as input (supervised training).

In this experiment, we developed a new ANN architecture designed to emulate the ecosystems processes driving NEE. This customized neural network (hereafter NN_C-part) is based on the concept that NEE, measured by the EC system, is the difference between RECO and GPP (Equation 1):

\[
\text{NEE} = \text{RECO} - \text{GPP},
\]

and that these two fluxes have individual drivers and dependencies. The uniqueness of our approach consists of imposing physical constraints in the proposed neural network, on the basis of the known properties of RECO and GPP. The overall structure is based on two subnetworks (Figure 1): one for retrieving RECO (sub neural network, SNN_RECO) and the other for retrieving GPP (SNN_GPP). The output of these two subnetworks \( S_{\text{RECO}} \) and \( S_{\text{GPP}} \) for SNN_RECO and SNN_GPP, respectively) is combined in the last neuron of the overall structure, where NEE is calculated and compared with the measurements in order to optimize the network’s weights (network training). The output signals of both subnetworks are constrained to be always positive, except for SNN_GPP, which is constrained to be 0 during nighttime, as photosynthesis requires light. However, RECO and GPP have opposite signs; therefore, the connection weight of RECO is fixed to be positive \( w_{\text{RECO}} = 1 \) and that of GPP to be negative \( w_{\text{GPP}} = -1 \), mirroring the sign convention of the measured NEE. Since the last transfer function is unbounded and linear, and the bias is fixed to 0, the last node effectively reproduces the equation of NEE as:

\[
\text{NEE}_{\text{ANN}} = R_{\text{RECO}} + G_{\text{GPP}}.
\]

where

\[
R_{\text{RECO}} = S_{\text{RECO}} * w_{\text{RECO}},
\]

and

\[
G_{\text{GPP}} = S_{\text{GPP}} * w_{\text{GPP}}.
\]

The two subnetworks have different input drivers but a similar structure including two hidden layers: the first hidden layer has n hidden neurons (different for the two fluxes) and a hyperbolic tangent sigmoid activation function (“Tansig”), which improved the performance of network optimization; the second has only one neuron and a log-sigmoidal activation function (“Logsig”) that constrains the output to positive values, allowing the implementation of Equation (2) in the last node. The SNN_GPP used to simulate GPP, has also a third layer where the measured incoming shortwave radiation enters as input in a node where a product is applied. This last node has a positive linear transfer function (“Poslin”); thus, the output from the SNN_GPP is 0 at nighttime and positive during the day. With this structure, the last node of the SNN_GPP mimics the light-use efficiency (LUE) approach: the output coming from the two previous hidden layers is an LUE that is then multiplied by the available shortwave radiation to estimate GPP. In this way, the drivers are used to estimate an instantaneous proxy of the LUE which takes into consideration seasonal and diurnal variability of carbon uptake, including photosynthesis saturation at high light, while shortwave radiation mainly defines the magnitude of...
the flux. This decreases the weight given to the incoming shortwave radiation in relation to other driving variables. Up to this step, the output of the two subnetworks \( S_{\text{RECO}} \) and \( S_{\text{GPP}} \) have positive signs; the sign convention of the NEE equation is finally restored in the last neuron of the overall structure by associating a positively signed weight to \( S_{\text{RECO}} \) and a negatively signed weight to \( S_{\text{GPP}} \).

### 2.2 Dataset used

For the purpose of this experiment, we used data from the FLUXNET2015 dataset (http://fluxnet.fluxdata.org) and, as an additional comparison, a synthetic dataset generated by the MUlti-layer Simulator of the Interactions between a vegetation Canopy and the Atmosphere (MuSICA) process-based model (Ogée, Brunet, Loustau, Berbigier, & Delzon, 2003, see section S4).

FLUXNET2015 includes meteorological and EC measurements that were quality checked and processed with standard tools (Papale et al., 2006; Pastorello et al., 2020) and provided with per-variable quality flags. For more details, see the webpage http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/.

Data used for training and validation of the neural network were taken from the “FULLSET” “TIER 1” collection. Among the sites available in this collection, we selected a subset of 36 study sites (listed in Table 1) on the basis of the data quality and data availability in order to ensure the best conditions for the partitioning methods comparison. Site-years were selected if the percentage of meteorological gap-filled data was less than 20\% and the measured NEE covered at least 10\% of both daytime and nighttime periods.

Among the variables stored in the FLUXNET2015 dataset, we used the GPP and RECO obtained with the two standard methods: the nighttime method (NT) from Reichstein et al., 2005, and the daytime method (DT) from Lasslop et al., 2010. In the NT method, the Arrhenius-type temperature–response curve of respiration as modeled by the Lloyd and Taylor equation (Lloyd & Taylor, 1994), driven by air temperature, is used for estimating RECO. This method makes use of the nighttime (when global incoming radiation <20 W/m\(^2\)) NEE observations (assumed as representative of RECO given the absence of photosynthesis at night) to fit the Lloyd and Taylor equation. There are two parameters to fit in the Lloyd and Taylor equation: the activation energy (\( E_0 \)) and the respiration rate at the reference temperature (\( R_{\text{ref}} \)). In the NT method, \( E_0 \) is estimated at an annual scale by calculating \( E_0 \) values every 15 days and then averaging the three with smaller uncertainty. Once \( E_0 \) is fixed, the \( R_{\text{ref}} \) parameter is estimated using short-term moving windows (4 days). On the basis of Equation (1), GPP is finally calculated by subtracting NEE from the daytime RECO extrapolated by the fitted model.

The DT method combines a rectangular hyperbola light response curve (Falge et al., 2001) for estimating GPP and the Lloyd and Taylor equation for estimating RECO (as in the NT method). In the

---

**FIGURE 1** The scheme of the customized neural network applied for retrieving gross primary production (GPP) and ecosystem respiration (RECO) from net ecosystem exchange (NEE) measurements. The two subnetworks (\( S_{\text{GPP}} \) left, in green, \( S_{\text{RECO}} \) right, in brown) are connected in the last step to estimate NEE. Inputs for RECO are air and soil temperature; soil water content; wind speed and wind direction; the day of the year (sine and cosine values of angular day of year); the average value of nighttime NEE. Input GPP are, in addition to the shortwave incoming radiation used in the product, air temperature; vapor pressure deficit; soil water content; potential and actual shortwave incoming radiation; wind speed and wind direction; a proxy of the mean seasonal GPP dynamic derived from the nighttime and daytime averaged NEE. See Section 2.3 of the main text for details.
DT method, nighttime NEE is used only to estimate the E0 parameter of the Lloyd and Taylor equation, while the remaining unknown parameters, including $R_{\text{ref}}$, are fitted using daytime measurements. In DT, the light response curve is driven by the incoming shortwave radiation (SW_IN), but it is adjusted for the effect of stomata closure due to the atmospheric evaporative demand using VPD as an additional driver of GPP (Lasslop et al., 2010). Similar to NT: (a) the Lloyd and Taylor model is driven by air temperature, (b) the model

<table>
<thead>
<tr>
<th>ID</th>
<th>Site code</th>
<th>IGBP</th>
<th>Lat</th>
<th>Lon</th>
<th>Validation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AU-Cpr</td>
<td>SAV</td>
<td>−34.00</td>
<td>140.59</td>
<td>*</td>
<td>Meyer, Kondríová, and Koerber (2015)</td>
</tr>
<tr>
<td>2</td>
<td>AU-DaP</td>
<td>GRA</td>
<td>−14.06</td>
<td>131.32</td>
<td></td>
<td>Beringer et al. (2011)</td>
</tr>
<tr>
<td>3</td>
<td>AU-Dry</td>
<td>SAV</td>
<td>−15.26</td>
<td>132.37</td>
<td></td>
<td>Cernusak, Hutley, Beringer, Holtum, and Turner (2011)</td>
</tr>
<tr>
<td>5</td>
<td>AU-Stp</td>
<td>GRA</td>
<td>−17.15</td>
<td>133.35</td>
<td>*</td>
<td>Beringer et al. (2011)</td>
</tr>
<tr>
<td>6</td>
<td>BE-Lon</td>
<td>CRO</td>
<td>50.55</td>
<td>4.75</td>
<td></td>
<td>Moureaux, Debacq, Bodson, Heinesch, and Aubinet (2006)</td>
</tr>
<tr>
<td>7</td>
<td>BE-Vie</td>
<td>MF</td>
<td>50.31</td>
<td>6.00</td>
<td></td>
<td>Aubinet et al. (2001)</td>
</tr>
<tr>
<td>8</td>
<td>CA-Qfo</td>
<td>ENF</td>
<td>49.69</td>
<td>−74.3</td>
<td>*</td>
<td>Giasson, Coursolle, and Margolis (2006)</td>
</tr>
<tr>
<td>9</td>
<td>DE-Geb</td>
<td>CRO</td>
<td>51.10</td>
<td>10.91</td>
<td>*</td>
<td>Anthoni et al. (2004)</td>
</tr>
<tr>
<td>10</td>
<td>DE-Gri</td>
<td>GRA</td>
<td>50.95</td>
<td>13.51</td>
<td>*</td>
<td>Prescher, Grünwald, and Bernhofer (2010)</td>
</tr>
<tr>
<td>11</td>
<td>DE-Kii</td>
<td>CRO</td>
<td>50.89</td>
<td>13.52</td>
<td>*</td>
<td>Prescher et al. (2010)</td>
</tr>
<tr>
<td>12</td>
<td>DE-Obe</td>
<td>ENF</td>
<td>50.79</td>
<td>13.72</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>DE-Tha</td>
<td>ENF</td>
<td>50.96</td>
<td>13.57</td>
<td>*</td>
<td>Grünwald and Bernhofer (2007)</td>
</tr>
<tr>
<td>14</td>
<td>DK-Sor</td>
<td>DBF</td>
<td>55.49</td>
<td>11.64</td>
<td>*</td>
<td>Pilegaard, Ibrom, Courtney, Hummelshøj, and Jensen (2011)</td>
</tr>
<tr>
<td>15</td>
<td>FI-Hyy</td>
<td>ENF</td>
<td>61.85</td>
<td>24.29</td>
<td>*</td>
<td>Suni et al. (2003)</td>
</tr>
<tr>
<td>16</td>
<td>FR-LBr</td>
<td>ENF</td>
<td>44.72</td>
<td>−0.77</td>
<td></td>
<td>Berbigier, Bonnefond, and Mellmann (2001)</td>
</tr>
<tr>
<td>17</td>
<td>GF-Guy</td>
<td>EBF</td>
<td>5.28</td>
<td>−52.92</td>
<td>*</td>
<td>Bonal et al. (2008)</td>
</tr>
<tr>
<td>18</td>
<td>IT-Bci</td>
<td>CRO</td>
<td>40.52</td>
<td>14.96</td>
<td></td>
<td>Vitale, Di Tommasi, D’Urso, and Magliulo (2016)</td>
</tr>
<tr>
<td>19</td>
<td>IT-Cp2</td>
<td>EBF</td>
<td>41.70</td>
<td>12.36</td>
<td>*</td>
<td>Fares, Savi, Muller, Matteucci, and Paoletti (2014)</td>
</tr>
<tr>
<td>20</td>
<td>IT-Cpz</td>
<td>EBF</td>
<td>41.71</td>
<td>12.38</td>
<td>*</td>
<td>Garbulsky, Penuelas, Papale, and Filella (2008)</td>
</tr>
<tr>
<td>21</td>
<td>IT-Mbo</td>
<td>GRA</td>
<td>46.01</td>
<td>11.05</td>
<td>*</td>
<td>Marcolla et al. (2011)</td>
</tr>
<tr>
<td>22</td>
<td>IT-Noe</td>
<td>CSH</td>
<td>40.61</td>
<td>8.15</td>
<td></td>
<td>Papale, Black, et al. (2015)</td>
</tr>
<tr>
<td>23</td>
<td>IT-Ro1</td>
<td>DBF</td>
<td>42.41</td>
<td>11.93</td>
<td></td>
<td>Rey et al. (2002)</td>
</tr>
<tr>
<td>24</td>
<td>IT-SRo</td>
<td>ENF</td>
<td>43.73</td>
<td>10.28</td>
<td>*</td>
<td>Chiesi et al. (2005)</td>
</tr>
<tr>
<td>25</td>
<td>NL-Loo</td>
<td>ENF</td>
<td>52.17</td>
<td>5.74</td>
<td></td>
<td>Moors (2012)</td>
</tr>
<tr>
<td>26</td>
<td>RU-Fyo</td>
<td>ENF</td>
<td>56.46</td>
<td>32.92</td>
<td>*</td>
<td>Kurbatova, Li, Varlagin, Xiao, and Vygodskaya (2008)</td>
</tr>
<tr>
<td>27</td>
<td>US-ARM</td>
<td>CRO</td>
<td>36.61</td>
<td>−97.49</td>
<td></td>
<td>Fischer, Billesbach, Riley, Berry, and Torn (2007)</td>
</tr>
<tr>
<td>29</td>
<td>US-MMS</td>
<td>DBF</td>
<td>39.32</td>
<td>−86.41</td>
<td>*</td>
<td>Dragoni et al. (2011)</td>
</tr>
<tr>
<td>30</td>
<td>US-NR1</td>
<td>ENF</td>
<td>40.03</td>
<td>−105.55</td>
<td>*</td>
<td>Monson et al. (2002)</td>
</tr>
<tr>
<td>32</td>
<td>US-SRM</td>
<td>WSA</td>
<td>31.82</td>
<td>−110.87</td>
<td>*</td>
<td>Scott, Jenerette, Potts, and Huxman (2009)</td>
</tr>
<tr>
<td>33</td>
<td>US-UMB</td>
<td>DBF</td>
<td>45.56</td>
<td>−84.71</td>
<td></td>
<td>Gough et al. (2013)</td>
</tr>
<tr>
<td>34</td>
<td>US-Whs</td>
<td>OSH</td>
<td>31.74</td>
<td>−110.05</td>
<td>*</td>
<td>Scott et al. (2015)</td>
</tr>
<tr>
<td>36</td>
<td>ZA-Kru</td>
<td>SAV</td>
<td>−25.02</td>
<td>31.50</td>
<td></td>
<td>Archibald et al. (2009)</td>
</tr>
</tbody>
</table>
parameters are fitted on short-term moving time windows to account for slow changing factors.

### 2.3 Input variables and data preparation

The variable used as target values in the $\text{NN}_{\text{C-part}}$ training was the half hourly $\text{NEE} (\mu \text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1})$ measured with the EC technique. In particular, we used the $\text{NEE}_{\text{CUT USTAR50}}$ variable (Pastorello et al., 2020). We used a comprehensive subset of micrometeorological variables measured at the flux towers as candidate drivers for $\text{NN}_{\text{C-part}}$, and additional variables derived from $\text{NEE}$, with the aim to explain variability of carbon fluxes at seasonal, daily, and half hourly resolutions. In particular, in terms of meteorological drivers, we used the $\text{SW\_IN} (\text{W/m}^2)$, $\text{VPD} (\text{kPa})$, air temperature ($\text{TA} [\text{°C}]$), soil temperature ($\text{TS} [\text{°C}]$), and soil water content ($\text{SWC} [%]$) in the upper soil layer (depth function of the site, see Pastorello et al., 2020), wind speed ($\text{WS} [\text{m/s}]$), and wind direction ($\text{WD} [\text{degrees}]$. Each one of these variables is involved to different degrees in carbon exchange processes or in $\text{NEE}$ measurements. $\text{SW\_IN}$ carries the photosynthetically active radiation that is a key variable for the light-dependent reactions of photosynthesis, while leaves regulate stomatal conductance in response to VPD (Lasslop et al., 2010); temperature has a key role in chemical reactions of biological processes; thus, it is involved both in photosynthesis and in the $\text{RECO}$ (Falge et al., 2001). $\text{SW\_IN}$ and VPD are used in the DT method to estimate photosynthesis while both DT and NT use TA to estimate $\text{RECO}$ only. Wind-related variables (WS and WD) affect the footprint of flux measurements (Arriga et al., 2017; Kljun, Calanca, Rotach, & W. and Schmid, H. P., 2015) which is important in particular if WD changes systematically (e.g., nighttime vs. daytime, morning vs. afternoon) and if the surrounding land cover is heterogeneous. The Julian day of the year (DOY) and the potential incoming radiation ($\text{SW\_IN\_POT} [\text{W/m}^2]$) were also used in order to provide information about seasonality and length of the day. The $\text{SW\_IN\_POT}$ was also aggregated at daily resolution to better track the seasonality of light conditions; the first derivatives of half hourly and daily $\text{SW\_IN\_POT}$ were also calculated to add specific information about the seasonal and diurnal dynamics of light (Bodesheim, Jung, Gans, Mahecha, & Reichstein, 2018; Papale, Black, et al., 2015; Papale, Migliavacca, et al., 2015).

Finally, since plant photosynthesis and $\text{RECO}$ change seasonally also due to substrate availability and management (in case of managed sites), daily average of nighttime $\text{NEE} (\text{NEE}_{\text{NIGHT}})$ and a proxy of $\text{GPP} (\text{GPP}_{\text{prox}})$ were calculated and used in input. In this case, gap-filled (Reichstein et al., 2005) half hourly $\text{NEE}$ was used in order to have a continuous time series. Proxy of $\text{GPP}$ was calculated by using $\text{NEE}_{\text{NIGHT}}$ and the average of daytime $\text{NEE} (\text{NEE}_{\text{DAY}})$ as follow:

$$\text{GPP}_{\text{prox}} = (\text{NEE}_{\text{DAY}} - \text{NEE}_{\text{NIGHT}}) \times k, \quad (5)$$

where $k$ is the fraction of daytime hours for each day.

Some of the drivers are periodic, which means that extreme values have similar meanings (e.g., the DOY 0 and 365 or the WD 0 and 360). To take this into consideration, we used their sine and cosine transformations instead of the original variable’s value.

The two subnetworks that estimate $\text{RECO}$ and $\text{GPP}$ use a different set of drivers, selected on the basis of the expected role and tests on performances and that are listed in Table 2. In particular, $\text{SNN}_{\text{RECO}}$ used TA, TS, SWC, WS, $\text{NEE}_{\text{NIGHT}}$, and sine and cosine of WD and DOY while the drivers for the $\text{SNN}_{\text{GPP}}$ were $\text{SW\_IN}$, $\text{SW\_IN\_POT}$ (half hourly and daily and their first derivatives), VPD, TA, SWC, WS, $\text{GPP}_{\text{prox}}$, and sine and cosine values of WD. Note that $\text{SW\_IN}$ is then used both as driver and as single input in the $\text{GPP}$ subnetwork for the product with the LUE (see Figure 1).

### 2.4 ANN training

Because our goal is to partition $\text{NEE}$ into its two fluxes components, the $\text{NN}_{\text{C-part}}$ was trained at the site level and year-by-year; we used only high quality measured half hourly $\text{NEE}$ (see Section 2.2) as training target. The training of the $\text{NN}_{\text{C-part}}$ was carried out using both nighttime and daytime measurements of $\text{NEE}$.

As commonly used in ANN training and application, inputs and outputs were normalized in the range $+1/-1$ by Equation (6):

$$X_{\text{norm}} = 2 \times \left( \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \right) - 0.5, \quad (6)$$

where $X_{\text{min}}$ and $X_{\text{max}}$ are estimated as the maximum of the absolute value of $X$ then negative signed in the case of $X_{\text{min}}$ and positive signed in the case of $X_{\text{max}}$; this preserved the zero value in both normalized

<table>
<thead>
<tr>
<th>Variable</th>
<th>Time resolution</th>
<th>GPP</th>
<th>$\text{RECO}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Measured</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{SW_IN}$</td>
<td>Half hourly</td>
<td>TA</td>
<td></td>
</tr>
<tr>
<td>$\text{SW_IN_POT}$</td>
<td></td>
<td>TS</td>
<td></td>
</tr>
<tr>
<td>VPD</td>
<td></td>
<td>SWC</td>
<td></td>
</tr>
<tr>
<td>TA</td>
<td></td>
<td>WS</td>
<td></td>
</tr>
<tr>
<td>SWC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Calculated</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{SW_IN_POT}$</td>
<td>Half hourly</td>
<td>Cos WD</td>
<td></td>
</tr>
<tr>
<td>First derivative</td>
<td></td>
<td>Cos WD</td>
<td>Sin WD</td>
</tr>
<tr>
<td>$\text{Sin WD}$</td>
<td></td>
<td>Sin(DOY)</td>
<td></td>
</tr>
<tr>
<td>Daily average</td>
<td></td>
<td>Cos(DOY)</td>
<td></td>
</tr>
<tr>
<td>$\text{SW_IN_POT}$</td>
<td>Daily average</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First derivative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{GPP}_{\text{prox}}$</td>
<td>Net ecosystem exchange night</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$SW\_IN is then used both as driver and as single input in the GPP subnetwork for the product with the LUE.
and original values without affecting the linearity of relationships between the original and the transformed time series.

Available records with all the inputs and the measured output were randomly split 25 times in sets for training (60% of the observations and, in general, at least 2,000 examples), test (20%) to avoid overfitting, and validation (20%) to assess the performances. This results in 25 datasets used for the ANN training. The ANN training used the mean square error as the cost function and the Levenberg-Marquardt backpropagation algorithm.

Five ANNs with different numbers of neurons in the first hidden layer of both subnetworks were trained: 10–13, 11–14, 12–15, 13–16, and 14–17 for SNN_{RECO} and SNN_{GPP} respectively. For each one of the 25 training datasets, the five networks with different structures were trained five times, changing the initial weights randomly, and then selecting the one with the best performance on the validation set. In case the sign of RECO and GPP did not respect the expected convention, the weights were reinitialized and the training was repeated. At the end of the process, there were 125 trained ANNs for each site-year (25 training datasets × 5 structures). To select the best among them, which were then used for the application, the ANN with the highest performances on the validation set was selected for each training dataset (obtaining 25 selected ANNs), and then, the five ANNs with the best model efficiency in predicting the overall NEE time series were finally selected and applied. RECO and GPP were calculated by averaging the output of the two subnetworks SNN_{GPP} and SNN_{RECO} from the five selected ANNs. The capacity of the ANN to simulate NEE was also analyzed (see Data S1).

### 2.5 Statistical analysis and evaluation

Results of this experiment were evaluated by comparing GPP and RECO from NN_{C-part} with the NT- and DT-based partitioning methods in the FLUXNET2015 collection. As complementary information, we repeated the same analysis for the gross CO_{2} fluxes derived from the synthetic NEE dataset provided by the MUSICA model (hereafter, we explicitly refer the synthetic CO_{2} fluxes modeled with MUSICA as GPP_{MUSICA}, RECO_{MUSICA} and NEE_{MUSICA}). The GPP and RECO estimates from the different methods were compared at half hourly, daily, and yearly time resolutions and were further evaluated about the consistency among methods for the seasonal cycle and its daily anomalies. The seasonal cycle was estimated by averaging the fluxes over 5 days while the daily anomalies were estimated by subtracting the seasonal cycle from each daily value. Finally, in order to compare the seasonal cycles across sites-years having different flux magnitudes, the annual average daily value was subtracted from each seasonal cycle. In addition to the seasonal cycles, the mean diurnal cycles were calculated and compared. To verify that the different drivers used in the NN_{C-part} had the role that is ecologically meaningful, functional relationships between micrometeorological drivers and retrieved gross CO_{2} fluxes were analyzed. In particular, the functional relationship between SW_IN and GPP, VPD and GPP, and TA and both GPP and RECO were considered and analyzed.

This was carried out in two different ways: (a) we simulated the flux responses in the NN_{C-part} varying artificially one driver at time, while the other drivers were kept fixed to the average condition of the season at midday; and (b) we compared the seasonal variations of gross CO_{2} fluxes with respect to the meteorological forces for both fluxes calculated using the NN_{C-part} results and those obtained using the NT and DT methods. The responses were obtained by averaging the values of 10 equal intervals step of the fluxes and drivers for 2 months.

The results of the machine learning approach were finally evaluated by looking at three specific responses of gross CO_{2} fluxes that are not explicitly accounted for by the DT and NT methods but can be indirectly considered in the NN_{C-part} method: (a) the instantaneous LUE responses to the diffuse to direct radiation ratio, (b) the hysteresis of the diel cycles of predicted RECO and GPP with respect to TA and SW_IN, and (c) the respiration pulse after a rapid increase in SWC occurring frequently in dry environments. Finally, the NN_{C-part} algorithm implemented here allows for an analysis of the role and importance of the different drivers (see Data S1).

### 3 RESULTS

#### 3.1 Cross consistency between NN_{C-part} and standard partitioning methods (DT and NT)

The half hourly GPP and RECO retrieved by NN_{C-part} were consistent with the output obtained by the NT and DT methods, with a slightly higher correlation (R^2 and lower RMSE) with the DT method for GPP estimation and with the NT method for RECO, in both cases higher than the correlation between DT and NT methods (Table 3). The mean squared error of partitioned gross CO_{2} fluxes between NN_{C-part} and the NT/DT methods was <1.53 µmol CO_{2} m^{-2} s^{-1} on average and comparable to the average value of the estimated random uncertainty of NEE ("NEE_CUT_USTAR50_RANDUNC," variable provided by FLUXNET 2015). Looking at the bias, the magnitude of both GPP and RECO retrieved by NN_{C-part} was slightly lower than the DT and NT methods, with the predicted values by NN_{C-part} closer to the DT method (though the spread among methods was very small). The average values of bias among methods were low on average (<27 g C m^{-2} year^{-1}) and comparatively lower with respect to the NEE random uncertainty reported in the FLUXNET2015 database (<10% of the reported "NEE_CUT_USTAR50_RANDUNC"). The agreement between NN_{C-part} and the two FLUXNET standard methods was significantly lower in the case of RECO suggesting that the differences occurred in the diurnal and seasonal cycles of RECO between NN_{C-part} and both DT and NT methods were comparatively larger than GPP.

High correlation and consistency among NN_{C-part} and the DT and NT methods in both GPP and RECO were found for daily, seasonal, and yearly values and also for the daily average anomalies with respect to the seasonal values (Figure 2). Daily, growing season, and yearly average values from NN_{C-part} were highly correlated with the two standard methods (R^2 > .83), in particular for GPP (R^2 > .97).
TABLE 3 Cross consistency among the retrieved gross primary production (GPP) and ecosystem respiration (RECO) by NN\textsubscript{C-part}, daytime method (DT), and nighttime method (NT) methods at half hourly time step, also after removing the mean daily value of the fluxes. Statistics reported in the table are the median of the study sites values, in brackets the 25th and 75th percentile.

<table>
<thead>
<tr>
<th>Gross CO\textsubscript{2} flux</th>
<th>Variability</th>
<th>Comparison</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>R\textsuperscript{2}</td>
</tr>
<tr>
<td>GPP</td>
<td>Overall</td>
<td>NN\textsubscript{C-part} versus DT</td>
<td>.96 (.93/.97)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NN\textsubscript{C-part} versus NT</td>
<td>.94 (.89/.95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT versus NT</td>
<td>.90 (.84/.93)</td>
</tr>
<tr>
<td></td>
<td>Removing mean daily value</td>
<td>NN\textsubscript{C-part} versus DT</td>
<td>.95 (.93/.97)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NN\textsubscript{C-part} versus NT</td>
<td>.90 (.86/.92)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT versus NT</td>
<td>.86 (.81/.89)</td>
</tr>
<tr>
<td>RECO</td>
<td>Overall</td>
<td>NN\textsubscript{C-part} versus DT</td>
<td>.73 (.54/.83)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NN\textsubscript{C-part} versus NT</td>
<td>.87 (.72/.92)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT versus NT</td>
<td>.78 (.65/.86)</td>
</tr>
<tr>
<td></td>
<td>Removing mean daily value</td>
<td>NN\textsubscript{C-part} versus DT</td>
<td>.21 (.09/.32)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NN\textsubscript{C-part} versus NT</td>
<td>.24 (.11/.38)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DT versus NT</td>
<td>.68 (.62/.76)</td>
</tr>
</tbody>
</table>

FIGURE 2 Scatter density plot showing the cross-consistency between partitioned gross primary production and ecosystem respiration by NN\textsubscript{C-part} (x-axis) and the ones by the daytime method and nighttime method methods (y-axis) aggregated at daily time step (a, e, i, m), yearly (b, f, j, n), and looking for the seasonal cycle (c, g, k, o) and daily anomalies (d, h, l, p).

In terms of bias, results by NN\textsubscript{C-part} were more consistent with the DT results confirming the findings from the subdaily comparison. High consistency was also found in the case of GPP daily anomalies (R\textsuperscript{2} > .88 and RMSE < 0.52 µmol CO\textsubscript{2} m\textsuperscript{-2} s\textsuperscript{-1}), but it decreased in the case of RECO (Figure 2).

This general consistency between NN\textsubscript{C-part} and standard methods (DT and NT) was confirmed by comparing the gross CO\textsubscript{2} fluxes derived by training the methods on synthetic NEE (NEE\textsubscript{MUSICA}, see section S4.1 in Data S1). The main findings from this comparison were that: (a) NN\textsubscript{C-part} and standard methods accurately retrieved both GPP\textsubscript{MUSICA} and RECO\textsubscript{MUSICA} at daily, yearly, seasonal timescales, and also the daily anomalies (see Figures S5 and S6); (b) NN\textsubscript{C-part} slightly outperformed the standard methods to estimate GPP\textsubscript{MUSICA} and RECO\textsubscript{MUSICA} particularly in the case of daily anomalies; (c) all methods slightly underestimated both GPP\textsubscript{MUSICA} and RECO\textsubscript{MUSICA} (bias ranging between −0.2 and −0.45 µmol CO\textsubscript{2} m\textsuperscript{-2} s\textsuperscript{-1}) with the largest discrepancy found in the case of DT method (for more details, see section S4.1, Figures S5 and S6).
3.2 | Consistency among the seasonal and diurnal dynamics of partitioned gross CO₂ fluxes

3.2.1 | Seasonal cycle

The seasonal pattern of the retrieved gross CO₂ fluxes by NNₐ part closely matched that of the DT and NT methods, with the only exception being that predicted RECO by NNₐ part was slightly lower in late spring/early summer (Figure 3, left). The predicted seasonal patterns were also similar among the represented International Geosphere–Biosphere Programme vegetation classes (Figure 3, right). Results showed a large degree of agreement among the three methods, in particular between NNₐ part and NT.

The seasonal dynamic of gross CO₂ fluxes derived from NEEMUSICA by NNₐ part, NT, and DT agreed consistently. However, all methods underestimated both GPP MUSICA and RECO MUSICA (in particular the DT method) during the growing season (between April and July, see section S4.2 and Figure S7 in Data S1).

3.2.2 | Mean diurnal cycle of GPP

We found a good agreement between the pattern of the mean diurnal cycle of GPP predicted by NNₐ part and FLUXNET standard methods (Figure 4). The GPP diurnal dynamic followed an asymmetrically bell-shaped curve with the maximum value reached before midday (Figure 4, top panels). Analyzing the differences more in depth, the mean diurnal cycle of GPP predicted by NNₐ part systematically diverged from the one predicted by the DT standard method, following a pattern that was consistent across the seasons (Figure 4, bottom panels). In particular, the GPP estimated by NNₐ part was lower during the first hours of the morning (just after the dawn) and in the late afternoon; instead, the GPP estimated by NNₐ part was slightly higher than DT from the morning until midday, where NNₐ part estimated higher GPP values. A similar pattern appeared when comparing the two standard methods (difference calculated as NT-DT) although shifted toward positive values, indicating higher values of GPP predicted by the NT method compared to the NNₐ part. The GPP derived from NEE MUSICA confirmed the differences we highlighted between DT and the other methods (see Section 4.3; Figure S8).

3.2.3 | Mean diurnal cycle of RECO

In comparison to GPP, the diurnal cycle of RECO modeled by NNₐ part showed larger differences with respect to the DT and NT methods, in particular regarding the magnitude of the flux. All the methods showed an increased respiration when moving toward midday, in particular during the growing season (Figure 5, top plots); nevertheless, the NNₐ part estimated values were lower than predictions from the other methods. Conversely, the differences between the two methods were smaller during the night. By analyzing the differences in detail, Figure 5 demonstrates that NNₐ part tends to underestimate RECO during the growing season, while the opposite occurs during the cold season. This pattern is consistent with previous studies that highlighted the importance of accounting for temperature-dependent respiration effects in carbon cycle modeling.
standard methods were negligible with a slightly higher RECO by NT in comparison to DT (Figure 5, lower plots, calculated as NT-DT).

This finding was not fully mirrored in the patterns of RECO\textsubscript{MUSICA} (see section S4.3 and Figure S9). In that case, NN\textsubscript{C-part} predicted higher RECO in comparison to both DT and NT, but more consistent with RECO\textsubscript{MUSICA}. Conversely, NN\textsubscript{C-part} predicted lower RECO than the NT method (and more consistent with DT) during the morning; the differences between DT and NT fitted on NEE\textsubscript{MUSICA} were also enhanced.
These differences increased during the nighttime hours, while the predicted RECO by the two standard methods became closer during the daytime hours (particularly when air temperature reached its maximum value). It is interesting to note that the peak of the maximum RECO by NN \textsubscript{C-part} was shifted compared to the DT and NT methods and closer to the one of RECO\textsubscript{MUSICA}, suggesting a better capacity to reproduce diurnal patterns when multiple drivers were involved (Figure S9).

3.3 | Functional relationships between partitioned fluxes and meteorological drivers

The functional response obtained by artificially varying the drivers in NN \textsubscript{C-part} was consistent with the current knowledge. The response of GPP to SW \textsubscript{IN} was as expected, with an increased photosynthesis due to the increased light that resulted as saturated at high radiation values (Figure 6a–d). The response to VPD clearly showed the effect of VPD limitation on photosynthesis (Körner, 1995) that was evident also at relatively low values of VPD (Figure 6e–h).

The response of photosynthesis to air temperature showed a strong and positive response at the start of the growing season that became flat in the summer period (Figure 6i–l). The functional relationships between RECO and TA followed the expected increase of RECO with TA (Figure 6m–p) except for the dry ecosystems (red points in Figure 6) where TA was not the main driving factor. These general patterns of relationships were consistent across the seasons and in the different vegetation types, despite the expected variability due to the ecosystem-specific properties.

When applied using the original measurements, the functional ecosystem response to the micrometeorological forces was largely consistent with the ones obtained using the DT and NT standard methods, although affected by multiple co-acting and confounding factors (Figure 7). The shape of the relations was slightly different in the four periods analyzed, but it was important to recall that these were average responses from sites in different climatic conditions. It could be noted that an almost perfect match was present in the case of GPP at this broad scale; in the case of RECO, the NN \textsubscript{C-part} showed in general lower respiration fluxes at high temperatures (so mainly daytime) as opposed to the other methods, although the pattern of response was basically the same. This could be an indication of a water resources limitation during the summer period and at higher temperatures, that could be detected by the NN \textsubscript{C-part} since it also used SWC as a driver. In fact, the functional relationships between RECO and SW retrieved by NN \textsubscript{C-part} highlighted a clear direct relationship between SWC and RECO in drought conditions, in particular for non-forested ecosystems (Figure S16).

3.4 | Additional ecological patterns simulated by NN \textsubscript{C-part}

3.4.1 | NN \textsubscript{C-part} instantaneous LUE and the ratio between diffuse to direct radiation

The instantaneous LUE estimated by NN \textsubscript{C-part} showed an increasing trend with respect to the increased diffuse to direct radiation ratio (here estimated by the proxy 1-SW \textsubscript{IN}/SW \textsubscript{IN,POT}; Figure 8a–d).

![FIGURE 6](image_url)
**Figure 7** Univariate functional ecosystem relationships between gross CO$_2$ fluxes retrieved by NN$_{C-part}$ (black lines) and FLUXNET standard methods (green and red lines for nighttime method and daytime method, respectively): gross primary production (GPP; µmol CO$_2$ m$^{-2}$ s$^{-1}$) and SW$_{IN}$ (W/m$^2$; a–d), GPP (µmol CO$_2$ m$^{-2}$ s$^{-1}$) and VPD (kPa; e–h), ecosystem respiration (µmol CO$_2$ m$^{-2}$ s$^{-1}$) and TA (°C; i–l). These patterns are derived from study sites located at the northern hemisphere (latitude > +15°N).

**Figure 8** Sensitivity of the light use efficiency (LUE) of NN$_{C-part}$ to the diffuse to direct radiation ratio. The trend of LUE by NN$_{C-part}$ with respect to the diffuse radiation ratio is reported in the panels (a–d; here LUE is normalized for comparison purpose). The difference between LUE (µmol CO$_2$/W) as estimated by NN$_{C-part}$ and that from the daytime method (DT) standard partitioning method is also reported (e–h); the statistics were aggregated per diffuse to direct radiation class (here estimated by the proxy 1-SW$_{IN}$/SW$_{POT}$ and reported as percentage). We reported the following statistics of x (x = LUE(NN$_{C-part}$) – LUE(DT)): the median values (bar), the mean value (*), the interquartile range (by brackets). Only LUE at midday is used in this figure.
When compared against the LUE estimated by the DT method, \( \text{NNC}_{\text{part}} \) showed a moderate but consistent higher sensitivity to the fraction of diffuse to direct light with a higher LUE when the diffuse to direct radiation ratio was higher than 40% (Figure 8e–h).

3.4.2 Hysteresis of the diel cycle of gross CO\(_2\) fluxes

The diurnal cycle of gross CO\(_2\) fluxes exhibited hysteresis, which we analyzed through the diurnal cycle of GPP with respect to SW\(_{\text{IN}}\) (Figure 9a–c) and the diurnal cycle of RECO with respect to TA (Figure 9e–g). We generally observed a clockwise cycle for GPP, with higher values of GPP in the morning in comparison to the afternoon. All the methods showed this pattern although it was more evident in the \( \text{NNC}_{\text{part}} \) and NT methods. We provided numerical estimates by calculating the integral area included in the hysteresis pattern (larger area indicates larger hysteresis) that was reported in Figure 9d.

For RECO, only \( \text{NNC}_{\text{part}} \) detected an appreciable hysteresis in the diurnal cycle (Figure 9h); this is expected because both the DT and NT methods use the same TA-dependent and invariant
relationships for the whole diel cycle. In general, the hysteresis of RECO is realized as a counterclockwise cycle, with the value of RECO, for the same TA, higher during the afternoon and night-time hours and lower during the morning. However, the counterclockwise cycle observed in the RECO patterns was not always observed, and in some cases, the opposite hysteresis cycle was found.

3.4.3 | RECO pulses due to SWC variations

The evaluation of the capacity to correctly reproduce rapid responses of RECO (respiration pulse) to changes in SWC (e.g., a post rain event after a prolonged drought period in arid environments) is in general difficult due to the sporadic nature of the events and the noise associated with the measurements. We evaluated the performances of the three models in a specific event that was clearly visible in the NEE time series (US-SRG site, Figure 10). The analysis of the predicted RECO shows that NN\textsubscript{C-part} was much more sensitive to the fast response of respiration pulses. The NT and DT methods, that did not have SWC as a direct input, showed a slow changing pattern with the peak of respiration significantly delayed compared to the measurements.

4 | DISCUSSION

4.1 | Consistency of NN\textsubscript{C-part} functional response with theory on plant physiology and with the DT and NT standard methods

The retrieved GPP and RECO as obtained from NN\textsubscript{C-part} were consistent with the estimates by the DT and NT methods both when trained directly on EC measurements and using synthetic data with NEE\textsubscript{MUSICA} (Figure 2; Figures S3 and S4). This was valid also for the relationship between micrometeorological drivers and gross CO\textsubscript{2} fluxes calculated by NN\textsubscript{C-part} and those implemented in the DT and NT partitioning methods (Figures 6 and 7).

The GPP response to light used by NN\textsubscript{C-part} is curvilinear and consistent with the pattern of that used in DT method (despite some small but systematic divergences). The pattern of the GPP–VPD relationship captured by NN\textsubscript{C-part} is consistent with current knowledge about the protective (physiological) mechanisms of stomata closure carried out by plants in response to the increase in atmospheric evaporative demands (Körner, 1995). There is also a general consistency of the GPP–VPD dynamics retrieved by NN\textsubscript{C-part} with those from the other methods (Figure 7) although the relationships include also the effect of other covarying factors, in particular incoming radiation. This stomatal regulation of photosynthesis is explicitly inserted in the formulation only in the DT approach. Interestingly, the NN\textsubscript{C-part} shows an effect of VPD also at very low values (Figures 6e–h and 7e–h), lower than the 10 hPa used as threshold in the DT method. The patterns of the TA–RECO relationships are also consistent across methods and in agreement with the expected trend due to the biochemical reactions involved in the organic matter respiration processes that find their optimal conditions at high (but not extreme) temperatures, if not limited by water availability.

The largest mismatch we found between NN\textsubscript{C-part} and the two standard methods of partitioning (NT and DT) is in late spring/early summer, with lower fluxes of daytime RECO predicted by NN\textsubscript{C-part} compared to both standard methods. When evaluated on the synthetic dataset, the mismatch between NN\textsubscript{C-part} and the standard methods was significantly lower except for the anomalies, which were better reproduced by NN\textsubscript{C-part}. In addition, the estimates from NN\textsubscript{C-part} better matched the diurnal dynamics of RECO\textsubscript{MUSICA} than the estimates from the standard methods (Figure S9); conversely the method that showed the highest mismatch with respect to RECO\textsubscript{MUSICA} was the DT method. All the methods underestimated both the reference GPP\textsubscript{MUSICA} and RECO\textsubscript{MUSICA} (despite the differences were very tiny) without any relevant effect on NEE. This could be related to a certain limitation of these methods to infer the complex relationships of the MUSICA model with the reduced set of drivers used in this experiment, but we have also to consider that both NN\textsubscript{C-part} and standard methods of partitioning were trained on noisy NEE signals (see section S4 in Data S1 for details) while both GPP\textsubscript{MUSICA} and RECO\textsubscript{MUSICA} used for reference are noise-free.

Recent papers highlighted possible limitations of the standard methods for partitioning stemming from the fact that they do not consider the inhibition of daytime leaf respiration (the Kok effect; Keenan et al., 2019; Wehr et al., 2016) that produces an overestimation of daytime RECO and GPP in the NT method (Wehr et al., 2016) and an underestimation of nighttime respiration (see Keenan et al., 2019) in DT. Although some of the differences between the RECO estimated by NN\textsubscript{C-part} and the other methods (see Figures 3 and 5 and Figure S9) are consistent with experimental data focused on this topic (e.g., Keenan et al., 2019; Wehr et al., 2016), it is not possible to demonstrate that the NN\textsubscript{C-part} method, as implemented in this experiment, is able to reproduce the light inhibition of leaf respiration. Moreover, the consistency between the DT and NT partitioned fluxes used as reference in this experiment, and the pattern emerged by other modeling experiences (see e.g., Jung et al., 2020, where DT and NT partitioned fluxes were globally upscaled) suggests that the Kok effect could have a minor relevance on biases of RECO and GPP estimates compared to other sources of NEE measurement uncertainties.

In terms of ecological responses, the NN\textsubscript{C-part} outputs showed a clear and interesting response to the diffuse/direct light with an increased LUE in diffuse conditions and with hysteresis of the diel cycle of GPP and RECO that are only partially shown by the standard methods. The latter showed also strong limitation in reproducing the respiration pulse while NN\textsubscript{C-part} showed this capacity, although tested only in one case. In summary, these findings highlight the strength of this approach that is able to reproduce these patterns even without being trained specifically for this.
In the case of RECO, both DT and NT used the same modeling approach, based on the Lloyd and Taylor model driven by the air temperature, resulting in high consistency between their results. However, it is known that soil temperature is also an important driver of RECO in particular to define its temporal (diurnal) dynamic (Lasslop et al., 2012; Wohlfahrt & Galvagno, 2017) because it is directly linked to soil respiration, which is a large contribution to the total RECO (Misson et al., 2007). The NNC-part method uses soil temperature (in addition to air temperature) as a direct driver of RECO and the analysis of the functional responses highlighted an important and direct effect (Figure S3). By looking at the analysis of RECO driver importance (see section S8 and Figure S20 in Data S1), it seems that TS and TA had, in general, a similar importance, but their relative weights changed by site. For this reason, it could be difficult to use both TA and TS in the same Lloyd and Taylor relationship without a priori information about the relative importance of the two variables. The multiple drivers approach implemented in NNC-part, where both TA and TS are used as drivers of RECO without any a priori assumptions also allows for the detection of hysteresis in the diurnal cycle of RECO (see Wohlfahrt & Galvagno, 2017 for details on the topic).

Another important driver of RECO that is not explicitly accounted for in the standard methods was SWC that is particularly important in dry sites (see section S8). The DT and NT methods use a dynamic parameter (R_w) estimated using moving windows to indirectly consider the slow changes in water availability (Reichstein et al., 2005). Soil water content, however, also has instantaneous effects in some ecosystems (e.g., the respiration pulses after short rain events, see Jarvis et al., 2007) that need the direct use of SWC as input, like in the NNC-part method, to be effectively modeled.

Other slow changing factors affecting the fluxes, such as phenological state, substrate availability, and management or other disturbance events, that are considered indirectly through R_w in the standard methods, are represented in the NNC-part approach through the use of averaged NEE-derived quantities as drivers.

Finally, the standard methods do not account for EC footprint variability, which can significantly affect the magnitude of the measured fluxes if the fetch is small and not fully homogeneous (see section S8). This could be considered, for example, by fitting the models per wind sectors and wind speed classes (or using footprint models), but the amount of available data could become critically low in less represented conditions. In the NNC-part, we used wind variables (direction, speed) as input in order to indirectly take into consideration footprint variability. Wind-related variables appeared to be important in a few study sites (see Figures S15, S16, and S17). By comparing the ANN trained by including/excluding wind variables, we detected an effect on performance (increase of RMSE), which highlights an important effect on the instantaneous estimated flux values. However, the role of wind variables is reduced in the yearly budget calculation of gross CO2 fluxes (~6 g C m⁻² year⁻¹, average value across study site) even though differences were detected site by site (site-specific differences ranged between -169 and +122 g C m⁻² year⁻¹). In summary, the
implemented machine learning approach gives the possibility to use variables as drivers which would be difficult to consider in process-based approaches, which require that the role and effects of each single variable being known and correctly encoded (at the ecosystem scale).

4.3 | Uncertainty and limitation of the proposed method

Despite the encouraging results, there are also limitations and uncertainty sources to consider when a pure machine learning approach is used. The absence of prescribed relationships and the few constraints fixed made the method robust against incorrect or incomplete formulations used in the standard methods but also led to higher sensitivity to the uncertainty in the data used in the training phase of machine learning, particularly when both NEE and the meteorological variables used as drivers are affected by long gaps. In the NN_{\text{C-part}}, the weights of the ANN are optimized against NEE; thus, the uncertainty of NEE could have a significant impact on their estimation. This could be particularly relevant for nighttime NEE measurements that can be affected by higher uncertainty (e.g., advection, large footprint) that, even if filtered, lead to long gaps. In these situations, a model based on ecological responses is in general more robust, and needs less data for a proper calibration than a data-driven approach.

Another limitation is related to the availability of the large set of input data that this approach requires in order to fully exploit its advantages, which are often not fully available at all FLUXNET sites. In this situation, the model can be used with a reduced set of drivers, but this could also lead to higher uncertainty and lower performance in specific sites, while models based on ecological processes like the NT and DT may be more robust because the response pattern is already prescribed.

5 | CONCLUSIONS

In this study, we propose a machine learning method for NEE partitioning (NN_{\text{C-part}}) using an ANN approach with a tailored structure to simultaneously retrieve GPP and RECO. From a methodological point of view, this approach is an example of simple hybrid modeling (Reichstein et al., 2019), combining a neural net with a "theoretical" ecosystem equation based on LUE.

The NN_{\text{C-part}} is designed to use as input a comprehensive set of drivers selected among the most commonly measured variables at EC sites. This property of NN_{\text{C-part}} (and of machine learning in general) can be exploited in future studies, for example, using as input other variables such as diffuse radiation, COS, 13C isotopes, or SIF. Additionally, NN_{\text{C-part}} does not make use of prescribed relationships between inputs and outputs and only a few constraints are introduced. Results in terms of the consistency of the gross CO\textsubscript{2} estimated fluxes by NN_{\text{C-part}} with the estimates from FLUXNET standard methods were good and the proposed model is able to reproduce magnitudes, patterns, and functional responses. However, the systematic differences emerged by the cross-consistency analysis and the ecological patterns captured by NN_{\text{C-part}} suggest that the implemented ANN reproduces fluxes dynamics more realistic than standard methods of partitioning. Being that said, further investigation are required in order to clarify the impact of the additional drivers used in NN_{\text{C-part}}, or possible missing functions and relationships in the standard methods.

Altogether, this new method provides GPP and RECO estimates that are based on purely data-driven empirical relationships without any assumption of the driver-output relations. This product would be of high interest for process model parameterization and validation as it avoids unwanted circularity, and for this reason, we propose it as a complement to the standard processing in large networks such ICOS, Ameriflux, or FLUXNET.

ACKNOWLEDGEMENTS

G.T. and D.P. conceived the experiment, G.T. implemented the method and did the analysis, J.O. provided the simulated data from MUSICA model, all the co-authors contributed in the preparation of the study. The authors thank Carlo Trotta for its help in the processing of the MUSICA data using the FLUXNET standard partitioning methods. This work used EC data acquired and shared by the FLUXNET community, including these networks: Ameriflux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia, and USCCC. The ERA-Interim reanalysis data are provided by ECMWF and processed by LSCE. The FLUXNET EC data processing and harmonization was carried out by the European Fluxes Database Cluster, Ameriflux Management Project, and Fluxdata project of FLUXNET, with the support of CDIAC and ICOS Ecosystem Thematic Center, and the OzFlux, ChinaFlux, and AsiaFlux offices. G.T. was supported by the projects H2020 BACI (grant agreement 640176) and by the European Research Council (ERC) under the ERC-2017-STG SENSIFLEX project (grant agreement 755617). J.V. would like to acknowledge the support from the European Research Council (ERC) under the ERC-2017-STG SENSIFLEX project (grant agreement 755617). G.C.V. would like to acknowledge the support from the European Research Council (ERC) under the ERC Consolidator Grant 2014 project SEDAL (grant agreement 647423). T.F.K. was supported by the NASA Terrestrial Ecology Program IDS Award NNH17AB86L. M. J. thanks the support of the e-shape project funded by the EU’s Horizon 2020 Programme (GA 820852). D.P. thanks for the support the RINGO project funded by the EU’s Horizon 2020 Programme (GA 703944) and the e-shape project funded by the EU’s Horizon 2020 Programme (GA 820852).

DATA AVAILABILITY STATEMENT

All the data used in the paper are included in the FLUXNET2015 collection available at https://fluxnet.fluxdata.org/ and shared under the CC4BY license. The codes for preprocessing and preparing data for partitioning, train the algorithm, select the best ANN, and produce the outputs are also shared under the BSD-Clause3 open license in GitHub at the address https://github.com/icos-etc/Partitioning_ANN.


