

1 The Regional MLP_{SGR}

This technical document provides instructions to derive the chlorophyll *a* concentration (*Chl-a*) by using the regional MLP_{SGR} for the Atlantic off the Southwest Coast of the Iberian Peninsula. The input quantity to compute *Chl-a* is the input normalized water leaving reflectance ρ_N at $\lambda = 490, 510$ and 560 nm. Section 1.1 to 1.3 describe the data pre-processing, the MLP_{SGR} function, and the data post-processing. The details to define the novelty index endorsing the validity of computed *Chl-a* values are in Section 1.4. Examples of algorithms coding in Matlab and Tables of coefficients derived from the field measurements at the Sagres site are provided on a case-by-case basis. The interested reader is addressed to the literature for additional details on the presented topics (see for instance [D’Alimonte et al., 2014] and references therein).

1.1 Data Pre-processing

The input normalized water leaving reflectance at the selected spectral wavelengths needs first to be log-transformed $l = \log(\rho_N)$, and then the z-score scaled as

$$\mathbf{x} = (\mathbf{1} - \boldsymbol{\mu}_l) / \boldsymbol{\sigma}_l \quad (1)$$

where $\boldsymbol{\mu}_l$ and $\boldsymbol{\sigma}_l$ are the mean and standard deviation of the entire set of log-transformed ρ_N values collected at the Sagres site. The Matlab code example for this processing is in Table 1, whereas the the $\boldsymbol{\mu}_l$ and $\boldsymbol{\sigma}_l$ coefficients are in Table 2.

Table 1: Listing of the Matlab code for the ρ_N data pre-processing.

```

% Pre-processing
logRho = log10(rho);
tmp = logRho - mu_l;
x = tmp ./ sigma_l;
```

Table 2: The set of parameters for the data pre-processing

$$\boldsymbol{\mu}_l = [-2.4266 \quad -2.4849 \quad -2.6089]$$

$$\boldsymbol{\sigma}_l = [0.1438 \quad 0.1259 \quad 0.1326]$$

1.2 The MLP_{SGR} Function

The pre-processed ρ_N values can then be ingested by the MLP_{SGR}. In summary, the MLP_{SGR} scheme consists of successive layers of units, with each unit in one layer connected to each unit of the next layer. Let $y = f(\mathbf{x})$ denote the MLP_{SGR} function, where the input vector \mathbf{x} has entries x_i for $i = 1, \dots, d$. The MLP_{SGR} hidden units values a_j are linear combinations of the input values:

$$a_j = \mathbf{b}_j^{(1)} + \sum_{i=1}^d \mathbf{w}_{ji}^{(1)} x_i, \quad (2)$$

where $w_{ji}^{(1)}$ represent the weights linking the input unit i to the hidden unit j , and $b_j^{(1)}$ is the bias adaptive parameter. The *activation* of the hidden unit j , identified with z_j and representing the input for the next layer, is obtained by applying the hyperbolic tangent activation function g to the summation of Equation 2:

$$z_j = g(a_j) = \frac{e^{a_j} - e^{-a_j}}{e^{a_j} + e^{-a_j}}. \quad (3)$$

The MLP_{SGR} output y is given by the weighted combination of the activations of the hidden units z_j through the second layer of weights $w_j^{(2)}$:

$$y = b^{(2)} + \sum_{j=1}^M w_j^{(2)} z_j, \quad (4)$$

where M indicates the number of hidden units and $b^{(2)}$ is the bias coefficient.

All these operations reduce to the few lines of Matlab code reported in Table 3. The set of corresponding MLP_{SGR} coefficient is in Table 4. Note that the MLP_{SGR} output has then to be post-processed as detailed in the next section.

Table 3: Listing of the code for computing the MLP_{SGR} output.

```

% Mlp computing
tmp = tanh(x * w1 + b1);
y = tmp * w2 + b2;
```

Table 4: Weight and bias parameters of the MLP_{SGR} function

$$\mathbf{w}^{(1)} = \begin{bmatrix} -0.5421 & 0.1274 & 1.3958 & -0.7542 & -0.4603 & 0.2405 & -0.1270 & -0.8025 & 0.6384 & 1.3734 \\ 0.1594 & -0.1563 & 1.5033 & 0.0787 & -0.5006 & -0.0986 & 1.0058 & -1.0930 & 0.0749 & -1.6038 \\ -0.6166 & 0.3531 & -1.5192 & -0.8465 & 0.1470 & 0.0548 & 0.5387 & -0.7408 & -0.6438 & 0.2567 \end{bmatrix}$$

$$\mathbf{b}^{(1)} = [-0.6816 \quad -0.1992 \quad 0.4899 \quad 0.1455 \quad 0.1777 \quad 0.2831 \quad 0.1228 \quad 0.4388 \quad 1.6791 \quad -0.9106]$$

$$(\mathbf{w}^{(2)})^T = [0.8607 \quad -0.0751 \quad -0.6629 \quad -0.1432 \quad -0.4240 \quad -0.3075 \quad 0.9862 \quad 0.7465 \quad -1.1217 \quad -2.1378] \text{ and } b^{(2)} = 0.0224$$

1.3 Data post-processing

Data post-processing is needed to transform the y value computed by the MLP_{SGR} into the final *Chl-a* value. The required mathematical operation is:

$$\text{alg1} = 10^{(y \cdot \sigma_c + \mu_c)}, \quad (5)$$

where σ_c and μ_c are the mean and the standard deviation of *Chl-a* values determined as the Sagres site. The code example and the corresponding coefficients are provided in Table 5 and 6, respectively.

Table 5: Listing of the code for post-processing the MLP_{SGR} output.

```
% Post-processing
tmp = y .* s_c;
chl_a = 10^(tmp + mu_c);
```

Table 6: Parameters for post-processing the MLP_{SGR} output.

$$\mu_c = -0.1879 \text{ and } \sigma_c = 0.4272$$

1.4 Applicability of the Regional MLP_{SGR}

The range of applicability of post-processed MLP_{SGR} results (that is, the computed *Chl-a* values) is estimated based on the novelty index η as follows:

$$\boldsymbol{\xi} = (\mathbf{x} - \boldsymbol{\mu}_c) \cdot \mathbf{A} \quad (6a)$$

$$\boldsymbol{\zeta} = \boldsymbol{\xi} / \sqrt{\boldsymbol{\gamma}} \quad (6b)$$

$$\eta = \|\boldsymbol{\zeta}\| \quad (6c)$$

where $\|\boldsymbol{\zeta}\|$ is the Euclidean norm of $\boldsymbol{\zeta}$, whereas the \mathbf{A} matrix and the $\boldsymbol{\gamma}$ array include respectively eigenvectors and eigenvalues from the Principal Component Analysis PCA of log-transformed ρ_N values measured at the Sagres. The Matlab code and the set parameters to compute the novelty index are in Table 7 and 8.

Table 7: Listing of the code to compute the novelty index.

```
% Novelty index
xi = (x - muIn) * A;
zeta = xi ./ sqrt(gam);
eta = sqrt(sum(zeta.^2));
```

Table 8: Parameters to compute the novelty index.

$$\mathbf{A} = \begin{bmatrix} 0.6481 & 0.6178 & 0.4454 \\ 0.5118 & 0.0798 & -0.8554 \\ 0.5640 & -0.7823 & 0.2645 \end{bmatrix}$$

$$\boldsymbol{\gamma} = [0.0412 \quad 0.0129 \quad 0.0001]$$

It is reminded that the estimated *Chl-a* shall be considered within the MLP_{SGR} applicability range when the corresponding η value is lower than 3 [D'Alimonte et al., 2014].

References

[D'Alimonte et al., 2014] D'Alimonte, D., Zibordi, G., Kajiyama, T., and Berthon, J.-F. (2014). Comparison between MERIS and regional high-level products in European seas. *Remote Sens. Environ.*, 140:378–395.