1 The Regional $\mathrm{MLP}_{\mathrm{SGR}}$ for API2 Retrieval

This technical document provides instructions to derive the the Algal Pigment Index 2 (API2) data product by using the regional $\mathrm{MLP}_{\mathrm{SGR}}$ 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 $560\,\mathrm{nm}$. Section 1.1 to 1.3 describe the data pre-processing, the $\mathrm{MLP}_{\mathrm{SGR}}$ 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 $\mathrm{l} = \log(\boldsymbol{\rho}_{\mathrm{N}})$, and then the z-score scaled as

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

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

$$\boldsymbol{\mu}_l = \begin{bmatrix} -2.4121 & -2.4698 & -2.5956 \end{bmatrix}$$

$$\boldsymbol{\sigma}_l = \begin{bmatrix} 0.1456 & 0.1328 & 0.1475 \end{bmatrix}$$

1.2 The $\mathrm{MLP}_{\mathrm{SGR}}$ Function

The pre-processed ρ_{N} values can then be ingested by the $\mathrm{MLP}_{\mathrm{SGR}}$. In summary, the $\mathrm{MLP}_{\mathrm{SGR}}$ 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 $\mathrm{MLP}_{\mathrm{SGR}}$ function, where the input vector $\mathbf{x}$ has entries x_i for $i = 1, \ldots, d$. The $\mathrm{MLP}_{\mathrm{SGR}}$ hidden units values a_j are linear combinations of the input values:

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

where $\mathrm{w}_{ji}^{(1)}$ represent the weights linking the input unit i to the hidden unit j, and $\mathrm{b}_j^{(1)}$ 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:

$$z_j \;=\; g\left(a_j\right) = \frac{e^{a_j} - e^{-a_j}}{e^{a_j} + e^{-a_j}}. \tag{3}$$

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

$$y = \mathrm{b}^{(2)} + \sum_{j=1}^{\mathrm{M}} \mathrm{w}_j^{(2)} z_j, \tag{4}$$

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

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

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

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

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

$$\mathbf{w}^{(1)} = \begin{bmatrix} -0.9549 & -0.4125 & -0.3321 & 0.1043 & -0.5055 & 1.0562 & -0.3186 & 0.7045 & 1.3208 & -0.6731 \\ 0.9401 & 0.2853 & -0.2978 & -0.0015 & -1.3192 & 1.3122 & 0.0058 & 0.4183 & 1.0452 & 0.6542 \\ -0.0922 & -0.4284 & 0.7543 & -0.0454 & -0.8413 & 1.1467 & 0.0849 & -0.0565 & 0.4818 & 0.0625 \end{bmatrix}$$

$$\mathbf{b}^{(1)} = \begin{bmatrix} 0.6307 & 0.4337 & -0.8282 & -0.1966 & 0.0243 & 0.7890 & -0.5071 & 0.1864 & -0.2228 & 0.7206 \end{bmatrix}$$

$$(\mathbf{w}^{(2)})^{\mathrm{T}} = \begin{bmatrix} 1.2481 & 0.3337 & 0.9707 & -0.7699 & -1.0198 & -0.6477 & 0.7931 & 0.6379 & -0.8338 & 0.6728 \end{bmatrix} \text{ and } b^{(2)} = \text{-0.3919}$$

1.3 Data post-processing

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

$$\mathrm{API2} = 10^{(y \cdot \sigma_c + \mu_C)}, \tag{5}$$

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;
chla = 10^(tmp + mu_c);
```

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

μ_c = -0.1478 and σ_c = 0.3310

1.4 Applicability of the Regional MLP_{SGR}

The range of applicability of post-processed MLP_{SGR} results is estimated based on the novelty index η as follows:

$$\boldsymbol{\xi} = (\mathbf{x} - \boldsymbol{\mu_c}) \cdot \mathbf{A} \tag{6a}$$

$$\boldsymbol{\zeta} = \boldsymbol{\xi}/\sqrt{\boldsymbol{\gamma}} \tag{6b}$$

$$\eta = \|\boldsymbol{\zeta}\|/n_\lambda \tag{6c}$$

where: 1) $\|\boldsymbol{\zeta}\|$ is the Euclidean norm of $\boldsymbol{\zeta}$; 2) 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; 3) n_λ is the number of selected wavelengths selected to compute data products.

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))./size(zeta, 2);
```

Table 8: Parameters to compute the novelty index.

$$\mathbf{A} = \begin{bmatrix} 0.5949 & 0.6041 & 0.5303 \\ 0.5796 & 0.1348 & -0.8037 \\ 0.5570 & -0.7855 & 0.2699 \end{bmatrix}$$

$$\boldsymbol{\gamma} = \begin{bmatrix} 0.0475 & 0.0130 & 0.0001 \end{bmatrix}$$

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 1 [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 Sensing of the Environment*, 140:378–395.