Stable isotope-based statistical tools as ecological indicator of pollution sources in Mediterranean transitional water ecosystems

Giovanna Jona-Lasinio\textsuperscript{a}, Maria Letizia Costantini\textsuperscript{b}, Edoardo Calizza\textsuperscript{b,*,} Alessio Pollice\textsuperscript{c}, Flavia Bentivoglio\textsuperscript{b}, Lucia Orlandi\textsuperscript{b}, Giulio Careddu\textsuperscript{b}, Loreto Rossi\textsuperscript{b}

\textsuperscript{a} Department of Statistical Sciences, Sapienza University of Rome, Rome, Italy
\textsuperscript{b} Department of Environmental Biology, Sapienza University of Rome, Rome, Italy
\textsuperscript{c} Department of Economical Sciences and Mathematical Methods, University of Bari, Aldo Moro, Bari, Italy

A R T I C L E   I N F O

Article history:
Received 9 October 2014
Received in revised form 3 March 2015
Accepted 9 March 2015

Keywords:
Isotope analysis
Bayesian modelling
Nitrogen pollution
Ecological indicator
Coastal ecosystems
Transitional waters

A B S T R A C T

N-stable isotope analysis of macroalgae has become a popular method for the monitoring of nitrogen pollution in aquatic ecosystems. Basing on changes in their \(\delta^{15}\)N, macroalgae have been successfully used as biological traps to intercept nitrogen inputs. As different nitrogen sources differ in their isotopic signature, this technique provides useful information on the origin of pollutants and their extension in the water body. However, isotopic fractionation potentially resulting from microbial nitrogen processing, and indirect isotopic variations due to effects of physicochemical conditions on algal nutrient uptake and metabolism, may affect anthropogenic N isotopic values during transportation and assimilation. This in turn can affect the observed isotopic signature in the algal tissue, inducing isotopic variations not related to the origin of assimilated nitrogen, representing a “background noise” in isotope-based water pollution studies.

In this study, we focused on three neighbouring coastal lakes (Caprolace, Fogliano and Sabaudia lakes) located south of Rome (Italy). Lakes were characterized by differences in terms of anthropogenic pressure (i.e. urbanization, cultivated crops, livestock grazing) and potential “background noise” levels (i.e. nutrient concentration, pH, microbial concentration). Our aim was to assess nitrogen isotopic variations in fragments of \textit{Ulva lactuca} specimens after 48 h of submersion to identify and locate the origins of nitrogen pollutants affecting each lake. \(\delta^{15}\)N were obtained for replicated specimens of \textit{U. lactuca} spatially distributed to cover the entire surface of each lake, previously collected from a benchmark, unpolluted site. In order to reduce the environmental background noise on isotopic observations, a Bayesian hierarchical model relating isotopic variation to environmental covariates and random spatial effects was used to describe and understand the distribution of isotopic signals in each lake.

Our procedure (i) allowed to remove background noise and confounding effects from the observed isotopic signals; (ii) allowed to detect “hidden” pollution sources that would not be detected when not accounting for the confounding effect of environmental background noise; (iii) produced maps of the three lakes providing a clear representation of the isotopic signal variation even where background noise was high. Maps were useful to locate nitrogen pollution sources, identify the origin of the dissolved nitrogen and quantify the extent of pollutants, showing localized organic pollution impacting Sabaudia and Fogliano, but not Caprolace. This method provided a clear characterization of both intra- and inter-lake anthropogenic pressure gradients, representing a powerful approach to the ecological indication and nitrogen pollution management in complex systems, as transitional waterbodies are.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Anthropogenic nitrogen loading of aquatic coastal environments significantly impacts ecosystem structure and functioning, leading to increased productivity, eutrophication, changes in microbial abundance and changes in both producer and consumer community composition (Duarte, 1995; Morand and Merceron, 2005; Rafaelli et al., 1998; Valiela, 1992; Valiela et al., 1997).
As coastal ecosystems, Mediterranean coastal lagoons are semi-closed and relatively small transitional water bodies, representing important but fragile habitats (Basset et al., 2006; Levin, 2001), which can be particularly affected by surrounding human activities. Direct and indirect anthropogenic nutrient loadings can arise from many activities, including treated and untreated sewage discharge, aquaculture, agriculture and livestock grazing, making it difficult to unambiguously determine the nutrient sources and their related impact on ecosystems (Grant et al., 1995; Jones et al., 2001; Orlandi et al., 2014). N-stable isotopes ($^{15}$N,$^{14}$N, expressed as $\delta^{15}$N) are increasingly used to investigate nutrient loading pathways in aquatic ecosystems (Costanzo et al., 2001; Dailer et al., 2010; di Lascio et al., 2013; Jones et al., 2001; Risk et al., 2009; Rožič et al., 2014). Atmospheric (natural) nitrogen can be distinguished from fertilizer and sewage-derived nitrogen on the basis of $\delta^{15}$N values (Dailer et al., 2010; Macko and Ostrom, 1994; Owens, 1987; Risk et al., 2009), with sewage-derived/organic N being characterized by higher $\delta^{15}$N values (around +6 to +38%) than inorganic/fertilizer-derived nitrogen (around -4 to +4%) (Dailer et al., 2010; Macko and Ostrom, 1994). The presence of different N inputs can thus be detected by changes in the isotopic signature of algal tissues in organisms transferred from an undisturbed location to the water body under study and allowed to uptake the dissolved allochthonous nitrogen (Costanzo et al., 2001, 2005; Jones et al., 2001; Orlandi et al., 2014).

Assuming that macroalgae reach isotopic equilibrium with the assimilated nitrogen source(s), human-derived nitrogen can be directly detected based on changes in the $\delta^{15}$N of macroalgae, and the allochthonous N input can be mapped by means of spatial analysis based on observed isotopic values (Costanzo et al., 2001). The isotopic values of nitrogen sources can be taken from the literature or measured at the origin of each N source potentially affecting the water body (Dailer et al., 2010). However, isotopic fractionation (i.e. discrimination between $^{15}$N and $^{14}$N during biochemical reactions) potentially resulting from ammonia volatilization (Heaton, 1986), nitrogen processing by bacteria (Leheman et al., 2003; Macko and Estep, 1984) and microalgal nutrient recycling in the water body (Pennock et al., 1996; Waser et al., 1998) may affect anthropogenic N isotopic values during transportation and processing (Dailer et al., 2010). Moreover, indirect isotopic effects can arise depending on nutrient concentrations (Pennock et al., 1996; Teichberg et al., 2008), nitrogen forms (Mariotti et al., 1981; Waser et al., 1998), temperature, pH and oxygen concentration, since these can all affect microbial activity and algal metabolism and nutrient uptake (Azov, 1982; Jones and Hood, 1980; Leheman et al., 2002; Rhee, 1978; Teichberg et al., 2007), representing background noise in isolate-based water pollution studies, i.e. inducing isotopic variations in algal tissue which are not related to the origin of assimilated nitrogen, potentially leading to erroneous interpretation of isotopic results. Such environmental background noise could be particularly influential in highly dynamic systems, as transitional aquatic environments and coastal lakes are. The relatively high degree of openness (Basset et al., 2006) of these environments makes them particularly sensitive to multiple and interacting stressors impacting the catchment area, with such elevated natural variability in both the environmental conditions and disturbance regimes potentially obscuring human-derived disturbance when occurring at low levels (Cloern, 2001; Vermeulen et al., 2011).

In this context, the spatial arrangement of point and non-point anthropogenic nutrient inputs, as well as their extent in the water body, may be particularly difficult to determine, complicating the discrimination between anthropogenic and natural variations in the measured physicochemical and biological parameters. In this case, modelling may be necessary in order to remove such noise from the observed isotopic values, allowing a reliable use of isotopic data to determine the origin of the dissolved nitrogen, the location of the allochthonous nitrogen inputs and their extent in the water body, with important implications for ecosystem management and conservation. Bayesian statistics is becoming increasingly popular in the ecological community. Using the words of Clark (2005) “advances in computational statistics provide a general framework for the high-dimensional models typically needed for ecological inference and prediction. Hierarchical Bayes represents a modelling structure with capacity to exploit diverse sources of information, to accommodate influences that are unknown (or unknowable) and to draw inference on large numbers of latent variables and parameters that describe complex relationships”. Then, the Bayesian paradigm allows us to merge theory with mechanistic understanding and empirical evidence, to assimilate diverse sources of information and to accommodate complexity often characterizing ecological datasets (Clark, 2006; Clark and Gelfand, 2006). It allows us to estimate the many contributions to the global uncertainty given by several sources, to model prior knowledge and to extract data information in a reliable way, representing a potentially useful approach with which complement isotopic field measurements by accounting for isotopic variations driven by environmental “noise” in pollution-monitoring studies. The Bayesian approach is not new in the statistical modelling of stable isotopes, having been successfully adopted in many protocols to determine animal diet using univariate and multivariate mixing models (Calizza et al., 2012; Erhardt and Bedrick, 2013; Hondula and Pace, 2014; Jackson et al., 2011; Phillips and Gregg, 2001; Richoux et al., 2014). Recently, the increasing interest in Bayesian statistics has led to the development of several tools for specific ecological analyses in the R environment (Jackson et al., 2011; Parnell et al., 2010) and to interesting debates, in particular, in the solving of under-determined stable isotope mixing systems, discussing whether adopt probabilistic approaches in a Bayesian framework (Semmens et al., 2013) or graphical approaches not allowing for probabilistic statements (Fry, 2013).

In the present study, we analysed isotopic changes occurring in a green macroalga, *Ulva lactuca*, deployed in three adjacent coastal lakes in central Italy that are known to differ both in anthropogenic pressure and in potential background noise, i.e. both abiotic and biotic factors potentially affecting isotopic results regardless of differences in anthropogenic nitrogen inputs. Based on environmental and isotopic data, we built a linear mixed-effects model (Zuur et al., 2009) to account for the uncertainty potentially affecting isotopic changes including background noise, potential measurement errors, replicate variability and spatial autocorrelation. The model was estimated in the Bayesian framework and model residuals (i.e. differences between observed and estimated isotopic changes at each sampling point within a given lake) were mapped and used for subsequent analyses. The novelty of this approach lies in its definition of a statistical model for stable isotopes that removes confounding factors from the recorded measures and allows for robust evaluation of the associated uncertainty. According to this, our aim was (i) to build an analysis protocol for isotopic data able to remove the effects of possible confounding factors and (ii) to apply this analysis protocol to the isotopic data in order to obtain reliable intra- and inter-lake comparisons and a robust description of the occurrence, spatial arrangement and extent of the anthropogenic nitrogen inputs affecting each lake.

2. Material and methods

2.1. Study areas

The three study areas were the lakes of Caprolace, Sabaudia and Fogliano (Latina, Italy), located on the Tyrrhenian coast of the Lazio region (42°28′00″ N, 12°51′00″ E) (Fig. 1). The maximum depths are
with respect to the mean size characterizing U. lactuca in the harvesting site. Three replicate algal specimens were deployed for 48 h at 50–60 cm water depth in plastic cages (one specimen per cage) at each sampling site. Algal deployment and collection was performed according to Orlandi et al. (2014). Before algal deployment, 1 cm² of the algal surface was dissected from each specimen in order to record the initial δ²¹⁵N value (δ²¹⁵N₀) at the individual level. After 48 h of submersion, each specimen was collected and transported to the laboratory. 48 h have been demonstrated to be a reliable exposure time when isotopic variations are quantified at the individual level in U. lactuca specimens (Orlandi et al., 2014). The isotopic variation (Δδ²¹⁵N) was calculated as the differences between the isotopic signature of each algal specimen after 48 h of submersion (δ²¹⁵N₁) and the isotopic signature of its own fragment dissected before deployment (δ²¹⁵N₀). This allowed us to calculate the isotopic variation at individual level while avoiding potential confounding effects arising from natural intraspecific variability in the isotopic signature. All samples were conserved at −80°C before isotopic analysis. At laboratory, samples were separately freeze-dried and ground to a fine homogeneous powder in a ball-mill (Fritsch, Pulverisette 23 with a zirconium oxide ball). C and N stable isotope analyses were performed by using an Elementar vario-MICRO CUBE analyser coupled with an Isoprime 100 mass spectrometer, operating as a continuous flow system. 2.0–2.5 mg dry-weight from each sample was analysed individually, both for algal fragments dissected before submersion and for specimens collected after submersion (Calizza et al., 2013; Orlandi et al., 2014). Isotopic ratios were expressed in δ units as the relative difference (in parts per thousand) between the sample and conventional standards (atmospheric N₂ (Air) for δ¹⁵N; PD-belemnite [PDB] carbonate for δ¹³C) in accordance with the formula δR(‰) = [(Rsample − Rstandard)/Rstandard] × 1000 (Ponsard and Arditì, 2000), where R is the heavy-to-light isotope ratio of the element (R = δ¹³C/δ¹²C or δ¹⁵N/δ¹⁴N). Results were monitored with reference to an internal standard calibrated to International Atomic Energy Agency reference materials (Caffeine: IAEA-CH₆). All samples were analysed twice. Measurement errors were found to be typically smaller than ±0.05‰.

2.3. Physicochemical variables

For each lake, physicochemical and microbiological data have been supplied by the Regional Agency for the Environmental Protection (ARPA), Provincial Section of Latina. Data were collected according to National standard sampling protocols, as requested by national laws (D.Lgs. 152/99, D.Lgs. 152/06, D.M. 260/2010, ISO 7899-2:2000, 2000, in APAT-IRSA-CNR, 2003). Physicochemical data were recorded during field samplings at a water depth between 40 and 60 cm, whereas microbiological counts were obtained at laboratory by filtering water samples on glass fibre (ISO 7899-2:2000 in APAT-IRSA-CNR, 2003). Such method allowed to count bacteria concentration (as the number of colony-forming bacteria on solid agar substrates) after incubation (36 ± 1°C for 48 h). Available data go back to 2006 until the time of this work and are the result of a monthly sampling (one sampling per month) which occurred in two different sampling locations per lake (near the centre and at the main outlet to the sea). Measured parameters showed a substantial stability in their mean, median, maximum and minimum values in the available time window. As it can occur in long-term datasets, data were affected by some quality issues such as missing values and unrecorded differences in measuring instruments. However, given the observed temporal stability of measured parameters, as we were interested in the main spatial tendency of these measurements (i.e. inter- and intra-lake differences), we reduced the impact of data quality issues by considering the overall time series median at each location within each lake. In order to

Fig. 1. Sampling area. The map shows the three study coastal lakes of Fogliano, Caprolace and Sabaudia (Latina, Italy) and the unpolluted site (Circeo, black circle) where algae were harvested few hours before their deployment within each study lake. The red dot in the small panel points the position of the sampling area in the Italian peninsula. Samplings occurred between May 14 and May 16, 2012. Lon: longitude, Lat: latitude. The black line marks the border of the urban settlement of the City of Sabaudia (19,287 inhabitants).

3.2 and 10 m for Caprolace, Fogliano and Sabaudia, respectively, and their surface areas are 2.3, 4 and 3.9 km², respectively. The three lakes are affected by varying sources and degrees of anthropogenic pressure (Fig. S1). Fogliano and Caprolace are SICs (sites of community importance) located in the Circeo National Park and are relatively well preserved. Sabaudia is privately owned, receives waste inputs from the homonymous town (19,287 inhabitants) situated on its landward side and is intensely exploited for fish and mussel farming. The seaward side of the lake has private land and low-density housing. A channel in the southern part represents the main connection with the sea. All the lakes are surrounded by cultivated crops. A livestock grazing area is located on the landward side of Lake Fogliano, which has two small water channels connecting the lake to the surrounding environment and the sea in the northern part, and a larger channel connecting the lake with the sea in the southern part. Caprolace presents two connections with the sea in the northern part, and it does not receive other water inputs from the surrounding area, thanks to a small water channel which surrounds and isolates the lake from adjacent crops and lands. Basing on lake surface and a regular sampling grid, sampling was performed at 24, 21 and 23 sampling sites at Caprolace, Fogliano and Sabaudia, respectively.

2.2. Field sampling and isotopic analyses

Algal specimens were collected from a benchmark, unpolluted site (Circeo) within the Circeo National Park (see Orlandi et al., 2014 for the description of this sampling location), few hours before their deployment in each study lake, which occurred between May 14 and May 16, 2012. Care was taken in the collection of specimens of very similar size, avoiding visibly smaller or bigger organisms
align the physicochemical and microbiological measurements with our sampling sites, we used the inverse distance weighted mean, considering the geographical distance from each sampling site and each measurement location within each lake. Environmental variables were found to be highly correlated with each other in several cases. After correlation analysis and ecological considerations on each study lake, we eventually kept three variables for modelling purposes: pH, nitric nitrogen (NN) and Enterococcus concentration (Ent). These were selected due to their potential effects on the observed algal Δ¹⁵N. Indeed, nitrification and denitrification by bacteria, coupled with ammonia volatilization, can affect the isotopic value of anthropogenic nitrogen (Dailer et al., 2010; Heaton, 1986; Lehman et al., 2002, 2003; Macko and Estep, 1984), while changes in N concentrations and water pH can affect both microbial activity and algal metabolism and nutrient uptake (Azov, 1982; Jones and Hood, 1980; Pennock et al., 1996; Rhe, 1978; Teichberg et al., 2007, 2008), with potential direct and indirect effects on the observed isotopic signatures of algal tissue. In addition, pH levels have been shown to vary both daily and seasonally in shallow coastal waters, showing relatively high values in shallow-water macroalgal habitats (Middelboe and Hansen, 2007), with direct effects on algal growth and photosynthesis, including species of the genus Ulva (Menéndez et al., 2001; Middelboe and Hansen, 2007; Nalewajko et al., 1997). Salinity was not considered for modelling purposes as (i) time series were characterized by several unrecorded changes in measuring instrument and protocols (both within and between lakes), suspect outliers and missing data and (ii) it was highly correlated with recorded values of NN, hence conveying no additional information in the model (Table S1).

2.4. The model

In order to provide an appropriate evaluation of isotopic signal variations in each lake, we needed to account for several sources of uncertainty such as the natural variability of physicochemical parameters, measurement errors, replicates variability and spatial autocorrelation if present. These sources may act in a similar way in all environments or be environment-specific. In order to account for this complex uncertainty, we modelled Δ¹⁵N as a function of the above-mentioned quantities and then used model residuals for subsequent spatial analyses. Model residuals were defined as the differences between the observed Δ¹⁵N value and the value estimated by the model (Fig. S2).

We used a linear mixed-effects model (Zuur et al., 2009) relating Δ¹⁵N to pH, NN and Ent, with spatially structured random effects and lake-dependent parameters. The model was specified hierarchically and was estimated in accordance with the Bayesian paradigm. The latter made it possible to describe uncertainty using prior probability distributions, assessing prior knowledge of the many sources of uncertainty and returning an entire probability distribution from which to draw inferences for each model component (Banerjee et al., 2004).

The model was formalized as follows: let \( Y_{ijr} \) denote the \( \Delta^{15}N \) at location \( i \) in lake \( j \) in replicate \( r \) and \( W_{ijr} \) be the spatially structured random effect at location \( i \) in lake \( j \). Given the random effect and the model parameters we then described \( Y_{ijr} \) as a Gaussian random variable:

\[
Y_{ijr} | W_{ijr}, \alpha_j, \beta_j \sim N(\alpha_j + X_{ijr} \beta_j + W_{ijr}, \sigma^2)
\]

where \( X_{ijr} \) is the vector of physicochemical variables evaluated at location \( i \) in lake \( j \) in replicate \( r \). The random effect \( W_{ijr} \) was assumed to be a zero mean spatial Gaussian process with exponential spatial correlation in Fogliano and Sabaudia, while it was a simple Gaussian random variable in Caprolace, where no significant spatial correlation was found. Prior distributions were chosen for each unknown model quantity: Gaussian distributions centred at zero with very large variance for the intercept and regression parameters, and diffuse truncated Gaussian distributions for the precisions \( \{1/\text{variances}, \text{mostly for computational reasons}\}

The priors for the decay parameters of the exponential spatial correlations, \( \varphi_p \) and \( \varphi_s \), were more informative. They were chosen based on prior knowledge of spatial isotopic variation in the Fogliano and Sabaudia lakes, as two uniform distributions over the intervals (0.65, 11) and (0.45, 9) for \( \varphi_p \) and \( \varphi_s \), respectively.

Model estimation required the use of Monte Carlo Markov chain (MCMC) algorithms to obtain samples from the posterior probability distributions of the model parameters. We implemented the model in JAGS (Plummer, 2003), considering two chains of 130,000 iterations with over-dispersed starts, keeping one in 10 samples from each chain and using the last 10,000 samples of each chain for inferences. Convergence was checked by visual inspection for MCMC traces and Gelman and Rubin’s convergence diagnostics (Gelman et al., 2003). Point estimates were computed using posterior means and interval estimation with 95% credibility intervals.

Once the model was estimated, the residuals were computed and used as a “clean” signal (i.e. free from the confounding effects of concomitant factors, hereafter referred as “post-processed” signal) as follows: let \( \hat{Y}_{ijr} \) be the value of the isotope variation predicted by the model at location \( i \) in lake \( j \) replicate \( r \), we then defined model residuals as: \( Z_{ijr} = Y_{ijr} - \hat{Y}_{ijr} \). These values were computed for each iteration of the simulation algorithm and were then used to produce point estimates of the post-processed signal at each replicate and sampling location and uncertainty evaluations given by the length of corresponding credibility intervals.

The final step in this analysis was the interpolation of the post-processed signal across the surface of each lake. We took the average over the replicates of the estimates of \( Z_{ijr} \) and interpolated these values using multilevel B-splines (Lee et al., 1997). The algorithm implementing multilevel B-splines takes a set of scattered data and produces tensor product B-spline surfaces in a computationally efficient way. It is suitable for points evenly distributed in space and it is implemented in the R software library MBA (mba.surf).

3. Results

3.1. Exploratory data analysis

Summary statistics for nitrogen isotope variation and the main physicochemical parameters are shown in Table 1. Median values of time-series data for abiotic covariates were spatially interpolated using the inverse distance weighted mean method. These initial explorations found the largest positive variations of \( \Delta^{15}N \) in Lake Sabaudia, while Fogliano had the highest level of nitric nitrogen (Table 1). Regarding remaining parameters, the three lakes showed a similar physicochemical structure.

Maps of the \( \Delta^{15}N \) produced using B-spline interpolation of observed data are shown in Fig. 2 using the same scale for the three lakes, no comparison among the lakes is possible without appropriate “standardisation”. In terms of intra-lake exploration, Fogliano and Sabaudia showed large homogeneous patches of isotopic variations, suggesting the presence of spatial correlation, confirmed by variogram analysis (Fig. S3). Caprolace showed no evidence of spatial correlation. From these considerations, we proceeded to build and estimate the model.

---

The Jags and R codes are available from the corresponding author upon request.

1 95% Credibility intervals are built using the 2.5% and 97.5% percentiles of the parameter posterior distribution as lower and upper bounds respectively.
2. Model selection and estimation

We compared the model given in Eq. (1) with two alternative specifications: 1. A model with different intercepts in each lake, the same regression coefficients for all lakes and the same covariance structure as the model in equation (1), i.e.

\[ Y_{ir} | \alpha_i, \beta, W_{ij} - N(\alpha_i + X\beta + w_{ij}, \sigma^2_Y) \]  

(2)

2. The same model presented in Eq. (2) with a spatial exponential correlation structure for Lake Caprolace.

Please notice that, in the current setting, physicochemical parameters vary only over space as they do not depend on the replicates. From a more formal point of view, this implies that X’s do not depend on the index r in the above equations.

The alternative models were compared using the deviance information criterion (DIC) (Spiegelhalter et al., 2002), the model in Eq. (2) returning the smallest DIC value (631.096). Table 2 shows point estimates and convergence diagnostics for the chosen model. The posterior mean is used as a point estimate, while 95% credibility intervals containing the zero value imply non significance of the corresponding parameter. The only intercept found to be significantly different from zero was Fogliano’s one, pointing to a distinctive behaviour of this lake, most likely due to the higher level of nitric nitrogen observed in the water body (Table 1): all remaining parameters were significant and highlighted direct relations between \( \Delta^{15}N \) and NN and Enterococcus, and an inverse relation between \( \Delta^{15}N \) and pH. The estimate of \( \sigma^2_Y \) measured the contribution of the sample replicates and unobserved factors to the overall variability. It was very close to 1, with a relatively small 95% credibility interval (0.74–1.15). Lake-specific components of the variability showed a large amount of uncertainty evaluated by 95% credibility intervals (Table 2): at Caprolace we had the smallest variation (\( \sigma^2 = 0.03 \)), followed by Fogliano (\( \sigma^2 = 0.04 \)) and Sabaudia (\( \sigma^2 = 0.16 \)). The size of the estimated spatial variances (\( \sigma^2_x, \sigma^2_y, \sigma^2_z \)) clearly showed a smaller order of magnitude with respect to the latent variability component measured by \( \sigma^2_Y \). The spatial covariance parameter estimates allowed us to evaluate the practical range2 as 0.5 km for Fogliano and 0.75 km for Sabaudia, suggesting highly localized effects.

3.3. Residuals computation and analysis

After model selection and fitting, we computed model residuals as described in Section 2.5. In order to analyse the distribution of these quantities in each lake, we computed the mean value of model residuals at each sampling location along with the credibility interval. As expected, in Caprolace, no spatial clustering was observed (Fig. 3), while in the other two lakes, uniform areas were found in terms of both the magnitude and the polarity (i.e. a negative or a positive variation) of the signal. Fogliano had three large positive values pointing to three possible pollution inputs (sampling sites: 1, 14, 23, Fig. S1), while Sabaudia was characterized by one major source of isotopic variation in the northeast part of the lake and a secondary source of variation on the seaward side. The uncertainty seemed larger in Fogliano, which was characterized by wider credibility intervals than the other two lakes. These results are shown and made easier to read in Fig. 3, which shows the interpolated surfaces of model residuals for each lake together with the maps of credibility interval bounds. The latters were obtained using multilevel B-splines applied to the bounds

---

Table 1 Summary statistics of isotopic variations and physicochemical measurements in Caprolace, Fogliano and Sabaudia coastal lakes (Latina, Italy).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Min.</th>
<th>Median</th>
<th>Mean</th>
<th>Max.</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caprolace</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Delta^{15}N )</td>
<td>-1.701</td>
<td>-0.014</td>
<td>0.074</td>
<td>2.003</td>
<td>0.396</td>
</tr>
<tr>
<td>NN</td>
<td>11.15</td>
<td>12.68</td>
<td>12.48</td>
<td>13.67</td>
<td>0.632</td>
</tr>
<tr>
<td>Enterococcus</td>
<td>0.187</td>
<td>1.236</td>
<td>1.445</td>
<td>2.845</td>
<td>0.665</td>
</tr>
<tr>
<td>pH</td>
<td>8.299</td>
<td>8.353</td>
<td>8.346</td>
<td>8.389</td>
<td>0.023</td>
</tr>
<tr>
<td>Fogliano</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Delta^{15}N )</td>
<td>-3.000</td>
<td>0.500</td>
<td>0.668</td>
<td>4.280</td>
<td>1.323</td>
</tr>
<tr>
<td>NN</td>
<td>34.05</td>
<td>34.86</td>
<td>34.77</td>
<td>35.27</td>
<td>0.359</td>
</tr>
<tr>
<td>Enterococcus</td>
<td>0.464</td>
<td>1.273</td>
<td>1.462</td>
<td>2.898</td>
<td>0.777</td>
</tr>
<tr>
<td>pH</td>
<td>8.355</td>
<td>8.363</td>
<td>8.365</td>
<td>8.379</td>
<td>0.008</td>
</tr>
<tr>
<td>Sabaudia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Delta^{15}N )</td>
<td>-0.529</td>
<td>3.451</td>
<td>3.478</td>
<td>7.102</td>
<td>0.846</td>
</tr>
<tr>
<td>NN</td>
<td>11.10</td>
<td>12.52</td>
<td>12.69</td>
<td>13.63</td>
<td>0.607</td>
</tr>
<tr>
<td>Enterococcus</td>
<td>0.231</td>
<td>0.981</td>
<td>1.226</td>
<td>2.895</td>
<td>0.702</td>
</tr>
<tr>
<td>pH</td>
<td>8.297</td>
<td>8.362</td>
<td>8.353</td>
<td>8.387</td>
<td>0.024</td>
</tr>
</tbody>
</table>

\( \Delta^{15}N \) refers to the differences between the final (i.e. after 48 h of submersion) and the starting (i.e. before submersion) isotopic signal of each algal specimen. NN: nitric nitrogen. The isotopic variation has been recorded during the sampling campaign occurred between May 14 and May 16, 2012. The physicochemical data were collected monthly each year starting from 2006 until the time of this study. Physicochemical data were recorded at two locations per lake: one at the main sea inlet, the other at the lake centre.

---

2 The practical range of the exponential covariance is defined as \( r = 3/\varphi \), where \( \varphi \) is the decay parameter of the exponential covariance function.
of credibility intervals obtained at each sampling location. These maps are easily comparable, confirming the above considerations and further highlighting hotspots, i.e. regions affected by nitrogen pollution sources (yellowish-red areas). Moreover, the maps of the credibility interval bounds illustrate the amount of uncertainty attached to hotspots: focusing on each lake, areas that cross the zero (from blush-green to white to yellowish-red in the colour scale) across the three maps are affected by a very large amount of uncertainty and the estimated values may not be statistically significant. As example, analysing the Sabaudia maps (cf. Figs. 2 and 3), a positive hot spot in the upper right side of the central panel was clearly detected. The extent of the area seemed to involve the entire upper part of the lake. However, using the change in the credibility interval maps, we were able to establish a more reliable dimension of the hotspot. The same rationale can be applied to negative hotspots, such as the blue area in the southern part of the Sabaudia lake.

### Discussion

Coastal lakes are complex environments, characterized by strong dependency on the surrounding terrestrial habitat (Basset et al., 2006) and, in this case, by their relatively small size and the complex spatial arrangement of potential anthropogenic stressors. This represented a very different starting point with respect to other isotope-based pollution studies carried out in marine coastal environments, where spatial dependency and pollutant dilution effects have been often detected on larger spatial scales (Costanzo et al., 2001, 2005; Jones et al., 2001; Orlandi et al., 2014; Careddu et al., 2015). In the same way, in our study, physicochemical and biological factors potentially affecting the isotopic signatures of both nitrogen inputs and algal tissue can plausibly be assumed to act on small spatial scales, producing intra- and inter-lake variations in the observed algal isotopic values. All this could lead to a possible environmental confounding effect affecting the marginal variation of isotopic outcomes following algal deployment in each lake. Nevertheless, the ability to deal with such environmental complexity is necessary in order to develop effective monitoring and conservation plans for the management of coastal lakes based on isotopic evidence of the origin, extent and location of nitrogen loadings.

We demonstrated that potentially confounding environmental effects on the ecological indication based on isotopic outputs returned by transplanted green macroalgae could be addressed by means of Bayesian statistical modelling, which made it possible to both “clean” the data by stripping the environmental noise and to consider the uncertainty associated to the obtained results. Better knowledge of factors causing possible confounding effects, obtained from long-term datasets referring to key physicochemical and biological parameters would make it possible to avoid potential biases arising from temporal outliers, which can affect all considered measurements in highly dynamic aquatic environments. Modelling confirmed the overall pollution gradient between lakes observed with the original isotopic data, and allowed to obtain lake-specific isoscapes unambiguously pointing out the location of the nitrogen pollution sources, the origin of the dissolved nitrogen and the extent of inputs in the water body. Indeed, based on “raw” (i.e. observed) δ15N algal isotopic signatures, Caprolace showed no point or diffuse anthropogenic nitrogen inputs and consequently the isotopic data showed no spatial dependency. Fogliano showed both intermediate anthropogenic pressure and spatial autocorrelation of isotopic signatures; the main source of organic nitrogen loading being detected in the southern part of the lake, corresponding to the main connection between the lake and the sea. Lastly, in Lake Sabaudia, we detected a diffuse isotopic enrichment throughout the lake, coupled with a strong spatial dependency of isotopic signatures. A major isotopic variation source was detected near the urban settlement of Sabaudia in the northeastern part of the lake, although the seaward side of the lake was generally characterized by higher isotopic increments than the landward side, consistently with the higher anthropogenic direct pressure (i.e. private houses and land) characterizing the seaward side (Dailer et al., 2010).

On the other hand, the proposed statistical modelling approach highlighted specific problems affecting each lake at the selected spatial scale that were not detected by the classical analysis of the system. Indeed, modelling made it possible to both (i) detect expected pollution sources in Fogliano Lake (i.e. the livestock grazing area on the landward side and the water channel flowing into the lake on the north side) which remained undetected basing on raw data and (ii) better characterize the extent and origin of pollution in Lake Sabaudia, where the expected predominant role of the urban settlement was highlighted by post-processed data. These results gave confidence on selected environmental parameters for isotopic data modelling as factors plausibly affecting observed isotopic results. Salinity (despite being not considered here for modelling given the poor quality of available records) was negatively correlated with values of dissolved nitrogen (NN). Besides modelling purposes, this suggest that NN inputs were directly associated to freshwater inputs, as it would be the case with inputs arising from the livestock grazing area in Fogliano or the runoff of urban loadings from the city of Sabaudia. The ability to associate uncertainty with the isotopic output by means of

### Table 2

Monte Carlo Markov Chain (MCMC) estimates and convergence diagnostics for the chosen model. αi (i = Caprolace, Fogliano, Sabaudia) is the model intercept, one for each lake.

| Parameter | Mean | 2.5%  | 25%  | 50%  | 75%  | 97.5%  | R² | n.eff
|-----------|------|-------|------|------|------|-------|----|------
| α⁰       | 19.01 | −20.67 | 5.29 | 18.95 | 32.64 | 59.13  | 1.00 | 8200
| α¹       | −62.06 | −96.48 | −73.70 | −62.17 | −50.14 | −27.59 | 1.00 | 5700
| α²       | 22.44 | −17.31 | 8.73  | 22.37 | 36.07 | 62.26  | 1.00 | 8100
| β₀       | 3.67  | 2.27  | 3.21  | 3.68  | 4.13  | 5.02   | 1.00 | 20000
| β₁       | 3.30  | 2.26  | 2.97  | 3.31  | 3.65  | 4.29   | 1.00 | 20000
| β₂       | −8.32 | −14.60 | −10.48 | −8.30 | −6.17 | −2.07  | 1.00 | 15000
| ϕ⁰       | 5.91  | 1.01  | 3.43  | 5.91  | 8.42  | 10.73  | 1.00 | 20000
| ϕ₁       | 4.04  | 0.60  | 1.80  | 3.64  | 6.12  | 8.68   | 1.00 | 20000
| σ₁²      | 0.03  | 0.01  | 0.02  | 0.03  | 0.05  | 0.11   | 1.00 | 13000
| σ₂²      | 0.04  | 0.01  | 0.03  | 0.05  | 0.09  | 0.31   | 1.00 | 20000
| σ₃²      | 0.16  | 0.03  | 0.16  | 0.29  | 0.44  | 0.95   | 1.00 | 8100
| σ₄²      | 0.91  | 0.74  | 0.84  | 0.91  | 0.98  | 1.15   | 1.00 | 20000

βi (i = Nitric), N[initrogen], Ent[erococcus], PH are the coefficients of physicochemical parameters, ϕi are the values of spatial correlation decay, σi² is the lake internal variability and σi₂² is the variance associated with the measurement error. The mean and the quantiles of the MCMC samples from the parameter posterior distributions are reported in columns. The last two columns show converge diagnostic of the MCMC algorithms. All converged with different degrees of chains autocorrelation. Convergence is achieved when the index is equal to 1. Autocorrelation between the samples is negligible when n_eff = n. Here n = 20,000.
Bayesian ecological modelling may represent an important contribution to ecosystem management, where decisions are supported by isotope-based pollution studies. Indeed, the consideration of different confidence interval limits associated with the mean isotopic variation calculated from post-processed data can be translated into different risk thresholds to be adopted by managers based on both the specific context (i.e. the specific water body and its priority in terms of management and conservation) and their ability to act.

Interestingly, this procedure (i) allowed us to remove background noise and confounding effects from the observed isotopic signals, (ii) produced maps of the three lakes providing a clear representation of isotopic signal variation even where background noise was high and (iii) increased the ability to unambiguously locate expected pollution sources and allowed to detect “hidden” pollution sources that would not be detected when not accounting for the environmental background noise. Thus, this method provided a clear characterization of both intra- and inter-lake anthropogenic pressure gradients, representing a powerful approach to the ecological indication and the management of nitrogen pollution in complex systems, as transitional water bodies are. In addition, the extensive and regular sampling grid adopted represented a powerful approach to describe not only the location, but also the path and the extension of nitrogen pollutants in each lake. Such approach seems to be particularly effective when an a priori spatial organization of potential disturbance sources could not be inferred from existing information (Dailer et al., 2012; Orlandi et al., 2014), providing a useful “scanning” tool to identify critical areas for future monitoring. The use of bioindicators (in comparison with chemical analysis performed on water samples) allowed to refer
observed results to the overall water volume filtered by macroalgae during the 48 h of submersion, with the further improvement of supplying evidences on the origin of nitrogen pollutants, with intuitive management implications. Undoubtedly, the isotopic analyses of water samples, despite their higher cost and being limited to the small water volume analysed, would produce important information on the isotopic baseline associated with dissolved nitrogen at the study-site scale. Lastly, the use of transplanted versus resident organisms allowed (i) to cover the entire surface of each lake according to a regular sampling grid and (ii) to directly measure post-deployment variations in algal isotopic signals, which, despite the potential of resident organisms to supply useful information on a longer exposure time to local nitrogen inputs [Vizzini and Mazzola, 2004], would have not been possible if considering autochthonous specimens. In particular, \textit{U. lactuca} was very scarce in the three coastal lakes at the time of this study, with few specimens highly varying in their size being observed in correspondence of lake shores, hence being not able to provide representative information on the spatial arrangement of nitrogen loadings at the lake scale over a coherent time period.

The possibility of implementing the model with different physicochemical and biological environmental parameters would allow us to focus on case-by-case specific environmental aspects, which are dependent on both the ecological problems affecting the considered habitat and operator's needs. The performance of the model was particularly satisfactory at the small spatial scale investigated, given the high environmental variability that can characterize transitional water systems [Basset et al., 2006]. The effectiveness of the model in relatively more homogenous and extended aquatic areas (e.g. coastal marine areas) remains to be tested. However, the proposed modelling approach is prone to implementation with environmental parameters potentially acting on larger spatial scales (i.e. from regional to global), such as temperature, CO$_2$ concentration, water acidification and seasonality. This will allow comparisons of isotope-based results from pollution studies carried out across a wide range of geographical, temporal and environmental conditions, representing a promising tool supporting a global perspective on isotope-based quantification of human impact on aquatic coastal environments.

Acknowledgments

We thank Dr Laura Aguzzi (ARPA) for the historical physicochemical and biological data and Dr. Pasquale Carlino for isotopic analyses. We also thank two anonymous referees for their helpful comments, which substantially improved the manuscript. This research was supported by Sapienza University grants 2012, PNRA-CONSIMA-L. Rossi 2010, Latina Provincial Administration 2011. The work was carried out in accordance with Italian laws.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolind.2015.03.006.

References


