CrossMark

ORIGINAL ARTICLE

Estimation of intrinsic aquifer vulnerability with index-overlay and statistical methods: the case of eastern Kopaida, central Greece

E. Tziritis[1] · L. Lombardo[2,3]

Abstract The intrinsic vulnerability of a karstic aquifer system in central Greece was jointly assessed with the use of a statistical approach and PI method, as a function of topography, protective cover effectiveness and the degree to which this cover is bypassed due to flow conditions. The input data for the index-overlay PI method were derived from field works and 71 boreholes of the area; the information was obtained, subsequently its critical factors were compiled which included lithology, fissuring and karstification of bedrock, soil characteristics, hydrology, hydrogeology, topography and vegetation. The aforementioned parameters were processed jointly with the aid of a GIS and yielded the final estimation of intrinsic aquifer vulnerability to contamination. Results were compared with an equivalent spatially distributed probability map obtained through a stochastic approach. The calibration and test phase of the latter relied on morphometric conditions derived by terrain analyses of a digital elevation model as well as on geology and land use from thematic maps. This procedure allowed taking into account the topographic influences with respect to a deep system such as the local karstic aquifer of eastern Kopaida basin. Finally, results were validated with ground truth nitrate values obtained from 41 groundwater samples, highlighted the spatial delineation of susceptible areas to contamination in both cases and provided a robust tool for regional planning actions and water resources management schemes.

Keywords Aquifer vulnerability · Groundwater · PI method · Kopaida plain

Introduction

Groundwater resources are often highly vulnerable to contamination from human activity, emerging the need for the appropriate protection measures. This point is highlighted in the EU Water Framework Directive (European Water Directive 2000). Karstic aquifers are well known for their particularly high vulnerability to contamination arising from their special characteristics, like thin soil covers, point recharge in dolines, shafts and shallow holes, as well as preferential flow paths in the epikarst and vadose zone. These characteristics result in contaminants easily reaching groundwater, where they are transported rapidly in karstic conduits over large distances (Zwahlen 2003). With respect to agricultural areas with intense activities that involve use of fertilizers and pesticides, groundwater is often threatened and should be appropriately managed (Giambelluca et al. 1996; Soutter and Musy 1998; Shaffer et al. 2001; Lake et al. 2003; Chae et al. 2004; Almasri 2008). Thus, defining integrated protection schemes for karstic groundwater systems is of vital importance for local communities and stakeholders to define and apply water management master plans, especially in environmentally sensitive regions.

The term 'vulnerability of groundwater to contamination' was first introduced by Margat (1968) and was based on the assumption that the physical environment provides

✉ E. Tziritis
  tziritis@gmail.com

[1] Hellenic Agricultural Organization, Soil and Water Resources Institute, 57400 Sindos, Greece

[2] Division of Physical Sciences and Engineering, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

[3] Institute of Geography, University of Tübingen, Ruemelinstr, 19-23 Tübingen, Germany

مدينة الملك عبدالعزيز
KACST للعلوم والتقنية

 Springer

to a degree natural protection to groundwater with regard to contaminants entering the subsurface environment. It is a relative, non-measurable and dimensionless property and can be distinguished into intrinsic vulnerability which accounts for the physical systems' characteristics (e.g., geology, soil, etc.) and specific which additionally takes into account the properties of a contaminant (Vrba and Zaporozec 1994).

Until now, several methods have been proposed for mapping the vulnerability of karstic aquifers; the most well know and frequently used include: EPIK (Dörfliger and Zwahlen 1998), REKS (Malik and Svasta 1999), RISKE (Pételet-Giraud et al. 2000) and RISKE 2 (Plagnes et al. 2005), PI (Goldscheider 2005), the Slovene approach (Ravbar and Goldscheider 2007), KARSTIC (Davis et al. 2002), COP (Vias et al. 2002), COP+K method (Andreo et al. 2009) and PaPRIKa (Kavouri et al. 2011). As a result, several researches worldwide have focused on assessing the vulnerability of diverse aquifer systems and can be found elsewhere (e.g., Prasad et al. 2011; Jayasekera et al. 2011; Kaliraj et al. 2014; Akpan et al. 2015; Kumar et al. 2015). However, results may differ significantly depending on the method applied and the study area. Therefore, the selection of the appropriate method chiefly depends on the specific site characteristics and availability of the required data. Conversely, very limited efforts have been generally made to support aquifer vulnerability studies using spatial predictive models, with the majority of these contributions focussed on factor weighting (Burkart et al. 1999; Nolan 2002; Worrall et al. 2002; Masetti et al. 2007; Mair and El-Kadi 2013; Nohegar and Riahi 2014; Pacheco et al. 2015). In this research, the reference model has been generated by adopting the protective cover–infiltration conditions (hereafter PI) method as its comprehensive application takes into account a broad spectrum of intrinsic causative factors (Goldscheider 2005). This result has been subsequently compared to a statistically derived model whose predictor space has been built to primarily represent superficial properties. The two different methodological approaches of index-overlay and statistical methods are combined and validate the final assessment of aquifer vulnerability.

The study area (Eastern Kopaida plain) is located in central Greece (Fig. 1). Kopaida plain is a highly cultivated region over the last decades; the eastern part which consists the study area of this research is characterized by intense farming activities and land use is dominated by agricultural land, apart from the hilly outcrops of the surrounding hydrological catchment. Despite the fact that Kopaida plain is a highly environmental vulnerable area, there are several scattered contamination sources, related to both point (e.g., manure from farms) or diffuse sources (fertilized plots).

Eastern Kopaida is also characterized by a unique karstic environment which is controlled by extended karstification phenomena (Tziritis 2009). These in general include a great number of sinkholes that enhance the rate of infiltration and subsequent water recharge and impose a significant potential threat to groundwater quality. The geological setting (Pagounis et al. 1994; Allen 1986; Tziritis 2008) (Fig. 1) at the bottom of the strata consists of a heterogeneous formation of Triassic dolostones and dolomitic limestones (G2); accordingly at the typical column follows a stratigraphical sequence of a tectonically driven metamorphic complex (mélange) of schists (F); this layer has no surface occurrence at the study area but occurs within the bedrock; locally due to intense tectonics F layer is absent and there is a direct stratigraphical contact between the lower (G2) and the upper formations (G1). Stratigraphy continues with a series of upper Jurassic bituminous limestones (G1). At their top is developed a paleo-karst surface filled by the chemically weathered material of the surrounding parent ultra-basic formations (C), composed by ophiolites and located only at the northern and northwestern part of the wider area. The upper bedrock sequence is completed with a highly karstified Cretaceous limestone (E) and the typical flysch (D). Post-Alpine formations at the spatial extension of the study area are composed of Quaternary fluvial, lacustrine and terrestrial deposits (A). However, Neogene formations (B) can be found in the wider region but do not exist in the typical column of research's study area.

The hydrogeological setting is controlled by the karstic network which drives the general groundwater flow from west to east; piezometric level varies locally from few meters to $\sim$160 m depending on substrate (Tziritis 2010). The main aquifer bodies in terms of water storage and yield are developed within the variable calcareous formations (dolostones and limestones) and are often karstified. Although typically can be distinguished three individual aquifers (e.g., cretaceous limestones, Jurassic limestones, and Triassic dolostones–limestones) their hydraulic connection creates a unified heterogeneous system, with variable permeabilities and degrees of karstification.

Methods and materials

Index-overlay method (PI)

The estimation of intrinsic aquifer vulnerability was conducted with the index-overlay PI method (Goldscheider et al. 2000). PI was selected due to its critical factors which were considered as more suitable for the specific conditions of the area (e.g., significant importance of protective cover,

Fig. 1 Geological map and typical stratigraphic column of the study area (adopted and modified by Tziritis 2008)
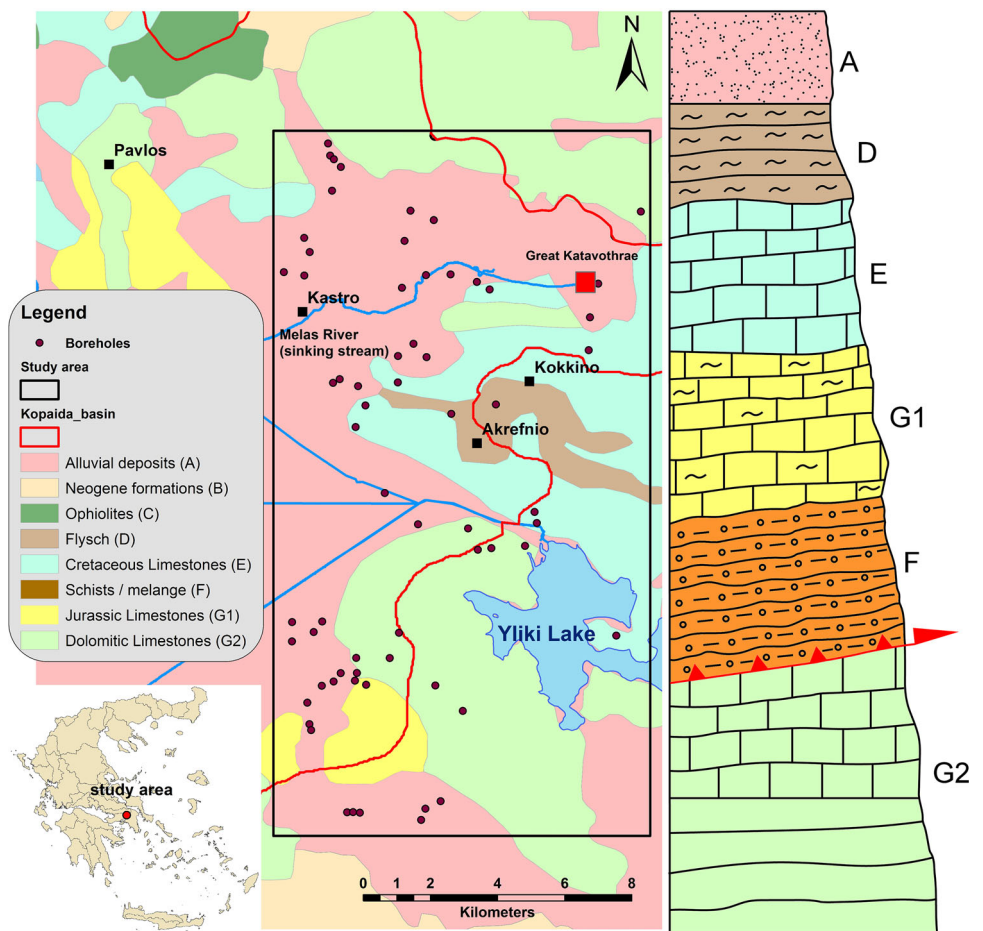
occurrence of "sinking stream", etc.). Input data were derived from various sources, including groundtruth values from 71 boreholes (Fig. 1) of the karstic aquifer as well as from field work and literature (Tziritis 2008). The parameters used included the following: vertical stratigraphy of the boreholes, piezometric level, aquifer's depth from the surface, fissuring and karstification degree of the substrate, topographic slope, lithology of the bedrock, type and origin of the quaternary deposits, occurrence of katavothres and sink holes, occurrence and route of blind river (a river sinking in a karstic formation) which in this case is River Melas, and finally spatial delineation of the hydrological sub-basins.

The PI method is a GIS-based approach with special consideration of karstic aquifers. It is based on an origin–pathway–target model (Fig. 2), where the origin is considered at the ground surface and the target is the uppermost aquifer. The pathway includes all the layers interfered between the extreme points of origin and target. The acronym stands for two factors, protective cover (P) and infiltration conditions (I). The simplified flowchart of the PI method is shown in Fig. 3.

Estimation of P-factor (protective cover)

The P-factor describes the protective function of the layers (soil, subsoil, non karstic rock, and karstic rock) between the ground surface and the water table, and is calculated according to a slightly modified version of the German (GLA) method (Holting et al. 1995) which is divided into five classes from 1 (very low protection) to 5 (very high protection). It expresses the impact of the protective cover on the basis of soils' effective field capacity (eFC), subsoil's grain size distribution, lithology, fissuring and karstification of the non-karstified and the karstified rock, thickness of all strata and mean annual recharge.

The above parameters were related to the "total protective function" (Goldscheider et al. 2000) to define the degree of total protection (P_{TS}):
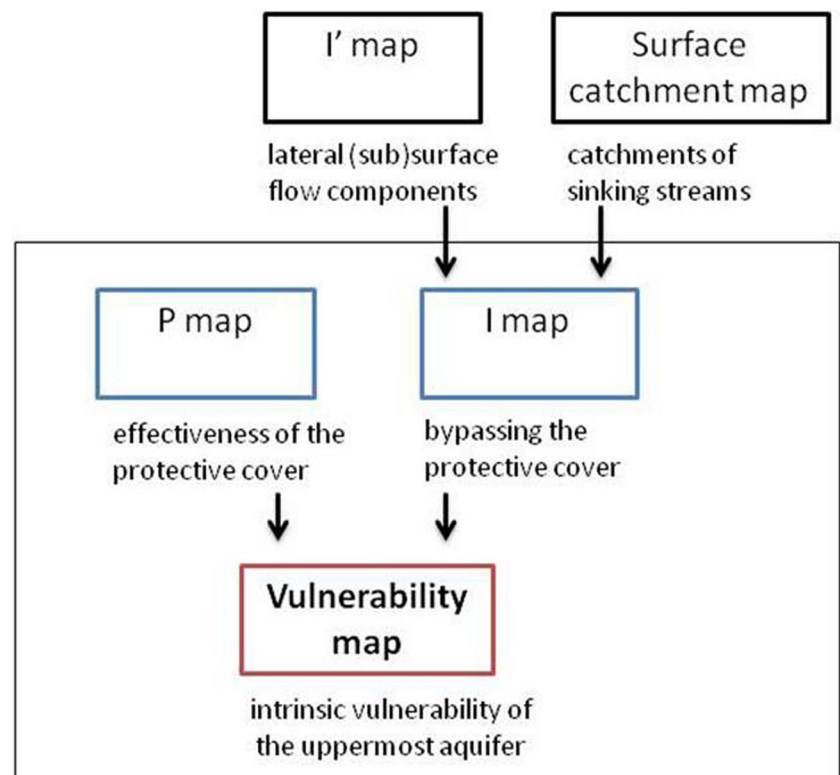
$$P_{TS} = \left[T + \left(\sum_{i=1}^{m} S_i \times M_i + \sum_{j=1}^{n} B_j \times M_j \right) \right] \times R + A \quad (1)$$

where T is the T-factor that refers to topsoil, S is the S-factor that refers to subsoil, B is the B-factor that refers to bedrock, M is the thickness (m) in the unsaturated zone,

Fig. 2 Illustration of PI method concept (after Zwahlen 2003)
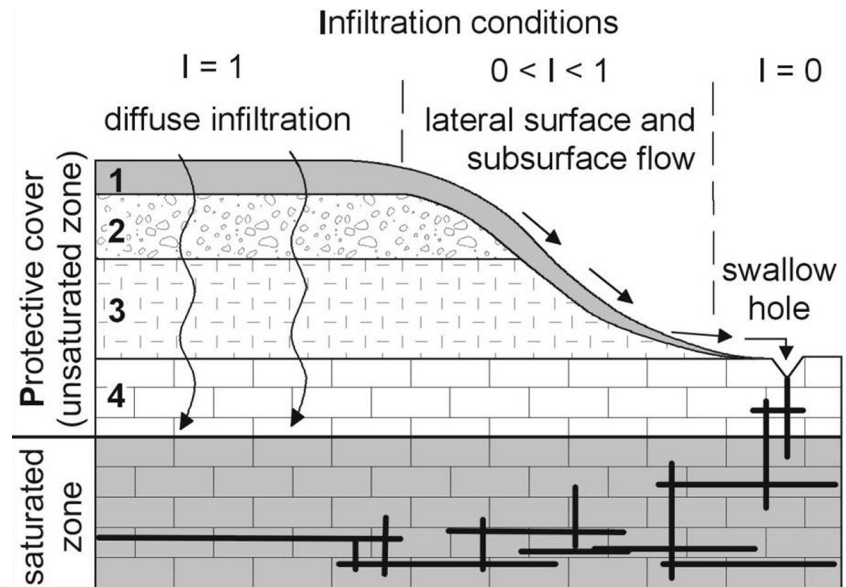
Fig. 3 Conceptual approach of the PI method (after Zwahlen 2003)

R is the R-factor that refers to recharge in terms of precipitation height, and A is the A-factor that refers to artesian conditions, if present.

The joint consideration of the above sub-factors resulted in the calculation of the final P-factor which is spatially distributed in Fig. 4. The range of P_{TS} values corresponds to a P-factor value (for more details see Goldscheider et al. 2000 and Goldscheider 2005) which ranges from $P = 1$ for

an extremely low degree of protection to $P = 5$ for very thick and protective overlying layers.

Estimation of the I-factor (infiltration)

The I-factor describes the infiltration conditions, and more specifically the bypass degree of protection cover, due to lateral and subsurface water flow in the catchment

Fig. 4 Spatial distribution of the P-factor (P-map), corresponding to the effectiveness of the protective cover in the study area
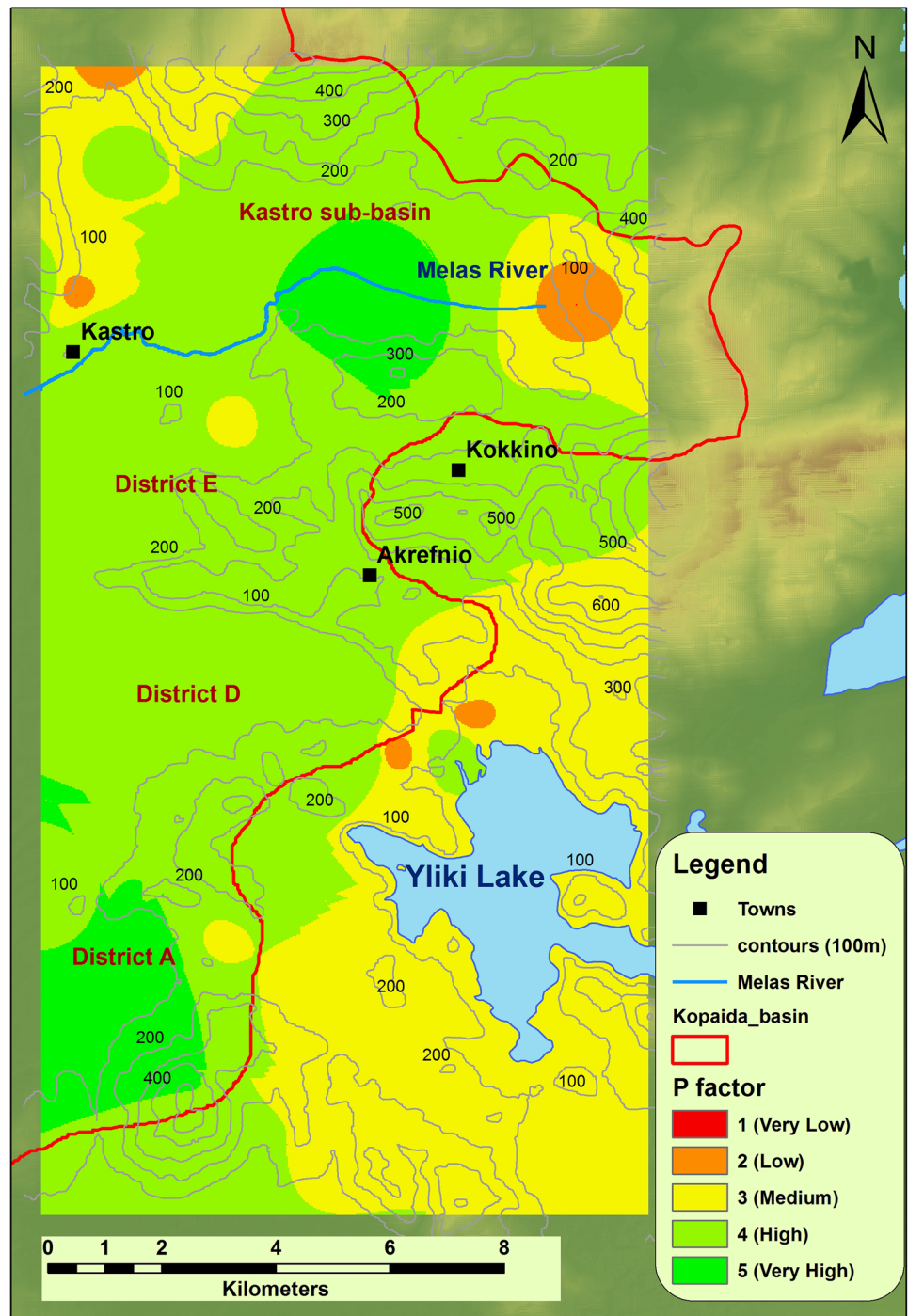
of karstic holes and sinking (blind) streams. If the protective cover is completely bypassed by a shallow hole through which surface water may pass directly into the karstic aquifer, then $I = 0$, while $I = 1$ if the infiltration occurs diffusely (e.g., on a flat, highly permeable and free draining surface). The intermediate values occur in catchment areas of variable slopes, depending on the proportion of lateral flow components. The final protection factor "π" is extracted by the combination of

P and I ($\pi = P \times I$) and subdivided into five classes, where $\pi \leq 1$ indicates very low degree of protection leading to extreme vulnerability, while $\pi = 5$ indicates a very high protection with subsequent very low vulnerability (Vrba and Zaporozec 1994; Goldscheider et al. 2000). The spatial distribution of "π" factor with the use of a GIS system defines the final outcome which is the assessment of the karstic aquifer's intrinsic vulnerability.

To determine the I-factor, three consecutive steps were carried out: firstly, the dominant flow process was estimated as a function of (1) topsoil's (or the uppermost layer) saturated hydraulic conductivity (m/s), and (2) the depth to low-permeability layers inside or below the topsoil (or the uppermost layer). The combination of the above produced a map (GIS coverage) which determined the final result of dominant flow process.

Subsequently followed the estimation of the I′-factor which depends on (1) results from the previous step (dominant flow process), (2) topographic slope, and (3) vegetation. The values (raster files) of each cell grid of the aforementioned parameters were added and finally yielded the I′ factor which ranged from 0.0 to 1.0. The result (GIS coverage) is called I′-map and shows the occurrence and intensity of lateral surface and subsurface flow.

Finally, the last step included the compilation of blind's river (sinking stream) catchment map. The conceptual approach of using this map is based on the assumption that lateral surface and subsurface flow may pose a risk to groundwater only if the contaminated water enters the karstic aquifer in a concentrated way, e.g., via a sinking stream, called as blind River; in the study area is Melas river (Fig. 5) sinks in a karstic hole at the northeastern region (great katavothre of Aghios Ioannis). According to the criteria imposed by the original PI method (Goldscheider et al. 2000), River Melas catchment was delineated in four buffer zones depending on relevant distances from the riverbed and the superficial occurrence of sinking holes and katavothraes. The obtained results included the construction of a relevant map (River Melas catchment map according to PI method criteria) with the aid of a GIS.

The final calculation of the I-factor embraced the combination of the previously assessed data and in turn compiled the I-map (Fig. 5) which shows the degree to which the protective cover is bypassed. Each grid cell is attributed a value derived from the intersection between I′-map (showing the occurrence and intensity of lateral flow) and River Melas catchment map (showing the sinking streams and their catchments).

Data processing and estimation of intrinsic aquifer vulnerability

The parameters/sub-factors which were considered to be spatially continuous (e.g., those derived by maps like soil, land use, surface geology, etc.) were input to GIS without any pre-processing as individual raster files, having a unique value for each grid cell. On the contrary, water table and thickness of subsoil were based on the spatial interpolation (IDW algorithm) of the data obtained from the 71 boreholes. In addition, the thickness of subsoil as well as the spatial distribution of subsurface stratigraphy

(including fissuring and karstification degree) was assessed with the aid of the derived geological map and further optimized with groundtruth data from the 71 boreholes. Hence, each cell was attributed a unique value which was derived by the combinational (joint) overlay of the all the considered parameters/sub-factors.

The final estimation of intrinsic aquifer vulnerability resulted from the combination of the individual P and I factors ($P \times I$) for the entire coverage of study area, divided into 25×25 m grid cells. The final outcome is the map of Fig. 6 which shows the spatial distribution of karstic aquifer's intrinsic vulnerability.

Statistical method

Despite the limited adoption of statistical methods for aquifer vulnerability studies, other branches of the earth sciences have extensively used these approaches in a wide range of applications (e.g., Conoscenti et al. 2013; Lombardo et al. 2014; Sahragard and Chahouki 2015). The procedure involves the identification of a dependent variable that a given algorithm, among the several available in literature, multivariately explains through a set of covariates. A calibration phase takes place initially when the selected algorithm learns to discriminate (through classification or regression) presence (in case of presence-only methods (Renner et al. 2015) or presence and absence [for presence–absence methods (Corani and Mignatti 2015)] of the dependent. A functional relation between predicted and predictors is subsequently derived and used to calculate probability of occurrences over broad areas such as catchments or regions. This prediction is then used to validate the model by testing its correctness against an unknown dataset. In this contribution, we adopted a presence-only approach known as maximum entropy (Phillips et al. 2006).

The dependent dataset has been generated by selecting among the 41 groundwater samples those with nitrate concentration above the screening value of 37 mg/L imposed by the Water Framework Directive (EU 2000). To increase the sample number a 73 m radius has been spanned from each borehole location defining as vulnerable all the centroids falling within these buffers in a synthetic 25-m grid coinciding with the study area. This operation has increased the vulnerable samples number to 258 points (virtual samples) spread across the eastern Kopaida plain but clustered in the proximity of the available boreholes. The predictor space coincides with the same 25-m grid being originally set by the available DEM. Primary and secondary topographic attributes have been calculated from the DEM obtaining: (1) elevation; (2) slope (Horn 1981); (3) profile curvature; (4) plan curvature; (5) catchment area; (6) topographic wetness

Fig. 5 Spatial distribution of the I-factor (I-map), corresponding to the bypass degree of the protective cover in the study area
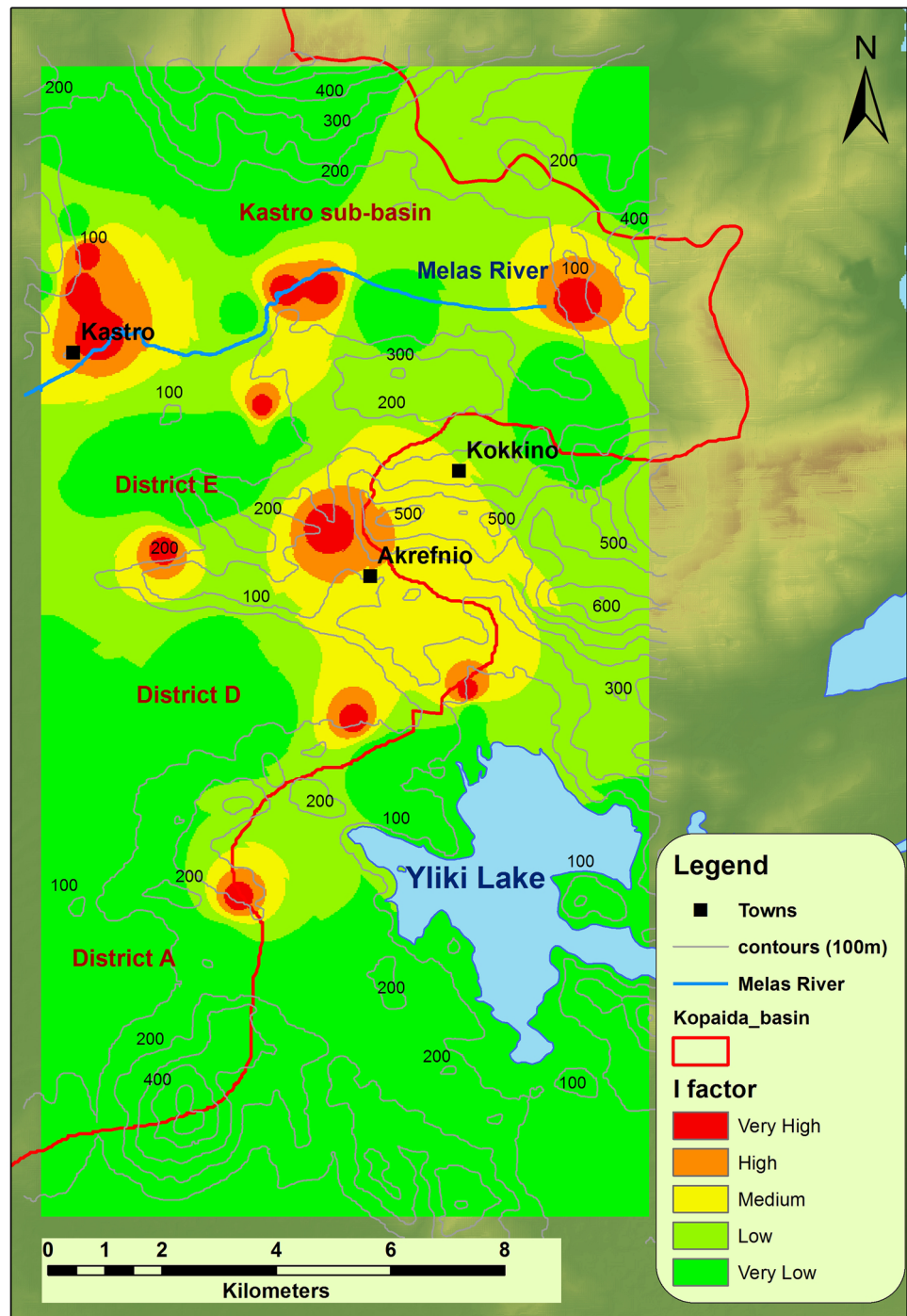
index (Beven and Kirkby 1979); (7) landform classification (Tagil and Jenness 2008). These predictors have been selected to represent superficial and shallow topographic influences on the aquifer vulnerability. Geology and land use have been also added as predictors being analogously gridded. The predictor types thus include both continuous and categorical classes, the latter being coded as shown in Table 1. Furthermore, the calibration phase has exploited the random 75 % of the 258 vulnerable points while the validation has been performed onto the remaining 25 %. Ultimately, the process has been repeated ten times to allow evaluating the stability of the prediction across the replicates. The performance-evaluation criterion has focussed on three main steps. The first one assessed the predictive skill of the models through area under the curve (hereafter, AUC; Phillips et al. 2006). This parameter represents the integer of a receiver operating characteristic curve (hereafter ROC; Parolo et al. 2008) which links

Fig. 6 Spatial distribution of karstic aquifer intrinsic vulnerability ("π" map)
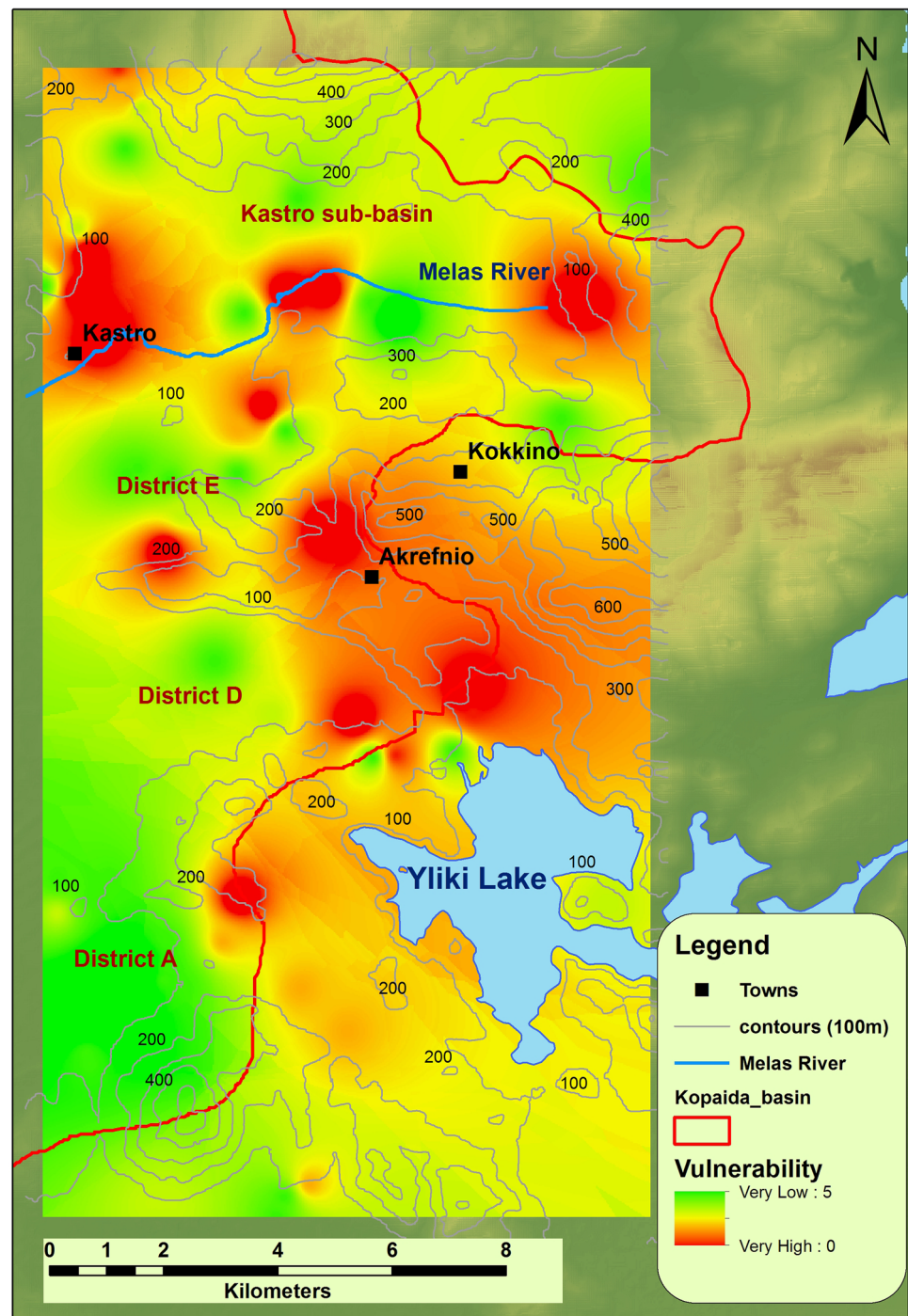
the proportion of positives cases that are correctly identified to the false-positive rate, for which Maxent applications are randomly extracted in the geographic space. Araújo and Guisan (2006) suggested AUC thresholds indicating average prediction between 0.7 and 0.8, good between 0.8 and 0.9 and excellent between 0.9 and 1. The second metric we adopted, investigated the percent contribution of each predictor with respect to the full model. This is commonly referred as predictor importance

(Lombardo et al. 2015) and its computation establishes how much a given predictor affects the final probability value at a given cell. Complementarily, the role of each predictor has been evaluated through response curves (Lombardo et al. 2015). These curves show the relation between the final probability and the domain of each predictor. By setting a probability threshold at 0.5 it is possible to distinguish anti-correlations and positive relations between the dependent and each predictor.

Table 1 Table linking the categorical codes to their real correspondent

Land use	
1	Discontinuous urban fabric
2	Industrial or commercial units
3	Road and rail networks and associated land
4	Mineral extraction sites
5	Non-irrigated arable land
6	Permanently irrigated land
7	Olive groves
8	Pastures
9	Complex cultivation patterns
10	Land principally occupied by agriculture, with significant areas of natural vegetation
11	Natural grasslands
12	Sclerophyllous vegetation
13	Transitional woodland-shrub
14	Water courses
15	Water bodies
Geology	
1	Alluvial deposits
2	Neogene formations
3	Upper jurassic limestones
4	Flysch
5	Upper cretaceous limestones
6	Triassic dolostones/limestones
7	Water bodies
Landform classification	
1	V-shaped river valleys
2	Local valley in plains
3	Upland incised drainages
4	U-shaped valleys
5	Broad flat areas
6	Broad open slopes
7	Flat ridge tops mesa tops
8	Local ridge/hilltops within broad valleys
9	Local ridge in plains
10	Mountain tops

Validation of results

The validation of an intrinsic vulnerability map is always a challenging and complex task that needs caution to avoid any erroneous interpretations and misleading results. The most common validation process is to measure the concentration of a surface-released potential contaminant and to compare the spatial distribution of its value with the vulnerability results. However, this process is not always straightforward and further caution is needed. For example, a common mistake is to ignore the interferences between the contaminant and the geological environment which may alter its concentrations and have a significant impact

on its overall fate and transport. The latter clearly highlights the importance of specific vulnerability which is by nature linked with those processes and is regarded as the most integrated approach for vulnerability assessment; however, even in cases of intrinsic vulnerability assessment, those interactions should be indirectly taken into account to validation process, to acquire representative results with minimized errors.

Another common error in vulnerability validation is the ignorance of lateral contaminant fluxes. By definition, aquifer vulnerability accounts for the vertical susceptibility of the system and is not considered for any lateral crossflows of contaminant plumes from adjacent hydrogeological units. Hence, practically validation with measured contaminant values at the saturated zone should be only performed at hydrologically "closed" systems, without any hydraulic connections with other units (surface or underground); if not, then potential migrations of contaminant plume(s) have to be taken into account prior to validation.

In the case of eastern Kopaida plain, validation was performed using the average nitrate values from 41 boreholes (Tziritis 2008) of wet (November–April) and dry (May–September) periods for the hydrological year 2004–2005. Essential precautions regarding the potential interactions and the migrating plumes were taken into account during validation process. Such interactions included abnormal deviations from the typical range of concentrations, for example, due to the development of strong anoxic zones within the heterogeneous karstic aquifer that dramatically decreased nitrates as a result of the reduction process. In addition, migrating nitrate plumes from adjacent basins often impart elevated values, which are not attributed to land use activities and physical properties of the specific study area.

The abnormal nitrate values due to redox conditions were excluded from the validation process as non-representative; their screening was based on previous hydrogeochemical assessments (Tziritis 2009, 2010) and it was further confirmed by their statistical population which were compiled by outliers, falling below the range of $m - 2s$ (where "m" is the mean statistical value and "s" is the standard deviation). Accordingly, the borehole samples which were proved to be affected by lateral crossflows and contaminant migration were excluded from the validation process too. The screening was performed on hydrogeological, hydrogeochemical and isotopic criteria from previously conducted researches (Tziritis 2009, 2010).

Results

The application of PI method resulted in the estimation of intrinsic aquifer vulnerability for the karstic system of eastern Kopaida plain. Intermediate steps, as already

described, included the estimation of P-factor (Fig. 4) which accounts for the degree of the protection cover effectiveness, and factor I (Fig. 5) which accounts for the degree of which this protective cover is bypassed due to flow conditions. Their joint assessment yields the intrinsic vulnerability of the karstic system expressed by "π" map (Fig. 6).

Based on the results of Figs. 4, 5 and 6, the regions with very low and low vulnerability are located in districts A′, D′ and E′, as well as in a part of Kastro sub-basin. The above regions are characterized by thick series of Quaternary deposits and very low topographic slopes, which are probably the key factors that control vulnerability. In addition, significant contribution to the overall result of low vulnerability at these areas should be given to the elevated values of soil effective field capacity (eFC) which are reflected at the high values of the topsoil sub-factor.

On the contrary, the higher values for vulnerability (high and very high) mainly appear as hot spots within the study area, except from the wider region around Akrefnio town which is spatially dominated by the outcrops of the karstic bedrock. Other key factors, mainly related to the aforementioned hot spots, include the absence (total or partial) of the protective cover, the increased degree of bypass flow, the elevated water table, and the high degree of karstification and fissuring of the bedrock (e.g., cretaceous and upper Jurassic limestones are by far more karstified and influenced by tectonic activity). The occurrence of karstified forms in surface is also a profound element for increased vulnerability values; for example, a local hotspot can be delineated at the end of Melas sinking stream (eastern Kastro sub-basin—great katavothrae) which directly enters the karstic network through a large sinkhole; in addition local hot spots are developed as a result of the great number of katavothraes at the fringes of hilly area between districts A, D and E.

Statistically derived vulnerability

The vulnerability across the Eastern Kopaida plain by means of statistical modeling has been assessed through the model prediction skill, predictor role and importance and spatial distribution of the probabilities. In terms of prediction skill the model produced excellent AUC values being on average 0.909. The corresponding ROCs are shown in Fig. 7.

Figure 8 highlights a prevalent influence of the elevation to the vulnerability, followed by the land use and plan curvature. For the present analysis, the average response curves of the nine adopted explanatory variables are shown in Fig. 9. The elevation presents a clear vulnerable trend between 90 and 110 m above sea level whilst the remaining

domain appears to play a negative role with respect to the vulnerability. As regards the land use role within the modeling procedure, the classes "road and rail networks and associated land", "permanently irrigated land" and "complex cultivation patterns" played a predominant role in assigning the vulnerable status to each cell in the study area. Negative plan curvature typically expresses surfaces sideward concave at a given cell thus affecting the convergence of the overland flow. The corresponding response curve shows a positive influence to the vulnerability with a contribution to the final probability decreasing as the curvature flattens, indicating linear water flows at zero. The positive or sideward convex morphology is shown to increase again and asymptote reached beneath 0.6.

Vulnerability map comparison

Comparing the maximum entropy (Fig. 10) to the PI-derived vulnerability maps has been done on an interpretative level as the former is expressed through a continuous probability estimation, whilst the second depicts the vulnerability through ordinal classes. In addition, maximum entropy has been performed only taking into account the topographic effects whilst the PI comprises shallow and deep parameters, being the most comprehensive method. This choice is due to the fact that spatial predictive models need for each of the considered predictors to cover the whole geographic space. To obtain this areal coverage even for deeper portions of the catchment, the only solution relies in interpolating from well data. However, as the well data itself is used as the dependent variable in the model, this will produce strong autocorrelation issues compromising the reliability of the output. As a consequence, differences arose between the two vulnerabilities. The most evident is expressed in terms of spatial variability. The latter is much greater for the statistically derived models with respect to the one showed for the PI. In light of these differences, the authors decided to evaluate and interpret the clear trends arising from the two maps.

The main vulnerable features have been captured by both methods. However, the main dissimilarity resides on the spatial extent of these features. In particular, the two models seem to agree within the areas of the District A and Kastro while over the sectors of Akfenio and Melas River the two maps substantially differ. The PI method in fact, highlights susceptible areas on a broader scale while the maximum entropy method emphasizes smaller features. Thus, the last two mentioned cases appear to be underestimated in terms of vulnerability through the statistic approach. This can be interpreted as an inherited weakness due to the difficulty of introducing predisposing factors linked to deep hydrogeological features.

Fig. 7 ROC curves calculated for each of the ten replicates, the training results are shown in *red* whilst the test results are shown in *blue*
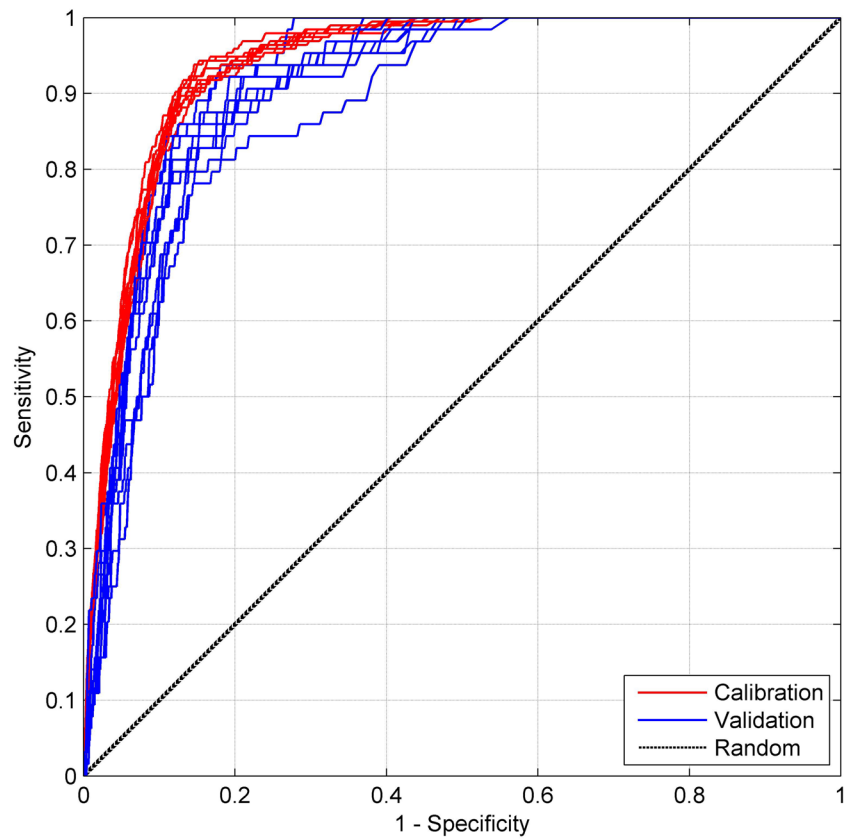
Fig. 8 Predictor importance plot showing the percentage of the contribution to the final probability for the selected covariates
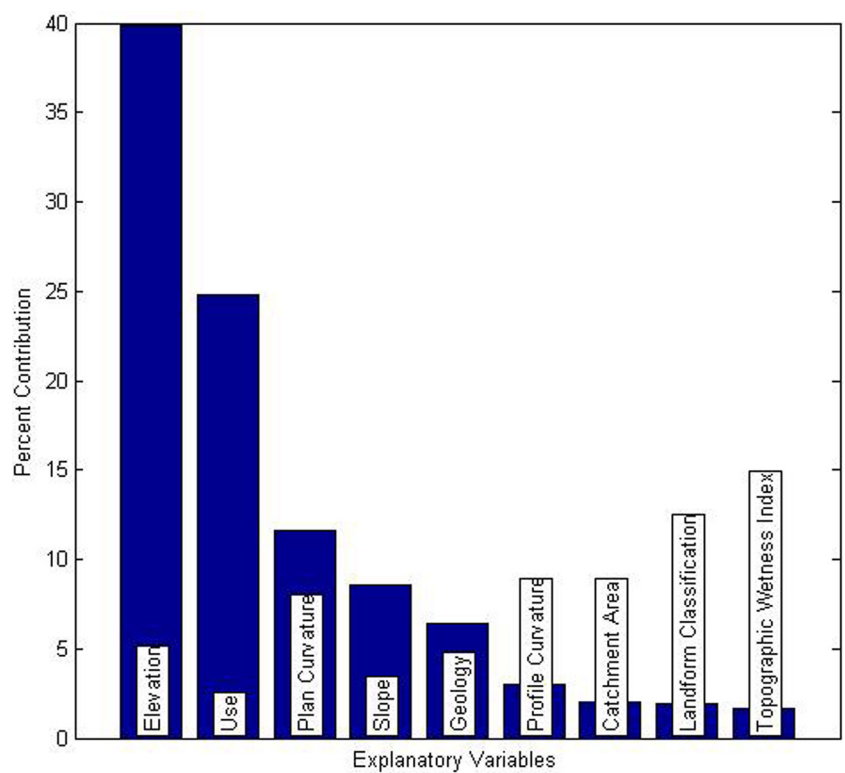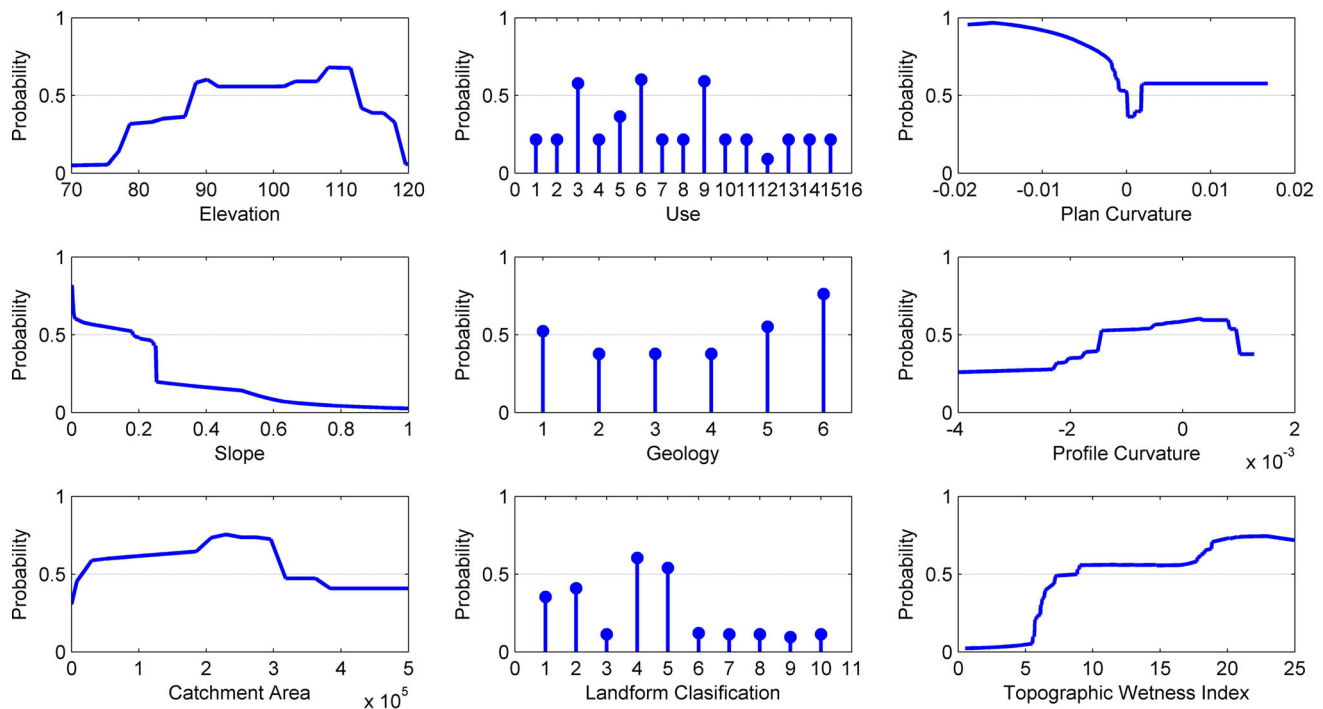
Fig. 9 Response *curves plot* shows the average relation between probability and the domain of each covariate

Discussion and conclusions

The intrinsic aquifer vulnerability was assessed with the use of geological data obtained from 71 borehole stations. The assessment was based upon PI method which is intended to estimate susceptibility of karstic aquifers. The approach was based on two main factors that describe the impact of the protective cover (P-factor) and the infiltration conditions (I-factor). The final assessment embraced all the critical parameters that influence aquifer vulnerability in the study area, like thickness and nature of geological formations, occurrence of karstification phenomena and topsoil. As a result, a map of intrinsic aquifer vulnerability was finally produced with the aid of spatial interpolation through a GIS system. Results revealed that higher vulnerability values were concentrated mainly as hot spots around surface karstified forms; additionally there were strongly linked with high piezometric level, reduced protection of the uppermost layers (topsoil, subsoil, bedrock) and high degree of karstification/fissuring of the vadose zone formations. Validation was performed with the use of groundtruth values of nitrates obtained from 41 groundwater samples located within the study area. Results revealed a good correlation between the real (measured) nitrate values and the vulnerability classes estimated by PI method. Complementarily, a presence dataset coinciding with the available nitrate values above a threshold of 33

(mg/L) has been built. An area of influence has been considered spanning 73 m from each of the aforementioned locations with high nitrates. The application of maximum entropy method allowed for predicting these vulnerable conditions through a set of topographic attributes integrated with geology and land use maps. Ten replicates have been produced, each time with a calibration proportion of 75 % of the initial presence dataset and the remaining 25 % used for validation. The results produced high-performance metrics indicating the models as reliable, despite the predictor set not including causative factors describing the deep hydrogeological settings. As a consequence the authors compared the two vulnerability maps anchoring the interpretation to the PI method and using the second method to infer topographic influences that could be imported to the reference PI model. The two maps agreed only at a broad scale while difference arose when looking at the details. In fact, the vulnerability depicted through the PI method is smoother and spatially homogeneous around the vulnerable features. Conversely, the statistically derived vulnerability rapidly varied across the area of interest from a pixel to another. This can be explained by taking into account how each model reflected the initial data used within the respective procedure. The map obtained by applying the PI methods reproduced geographic features influenced by the borehole distribution and as such affected by the borehole density across a given

Fig. 10 Intrinsic vulnerability map obtained by applying maximum entropy modeling
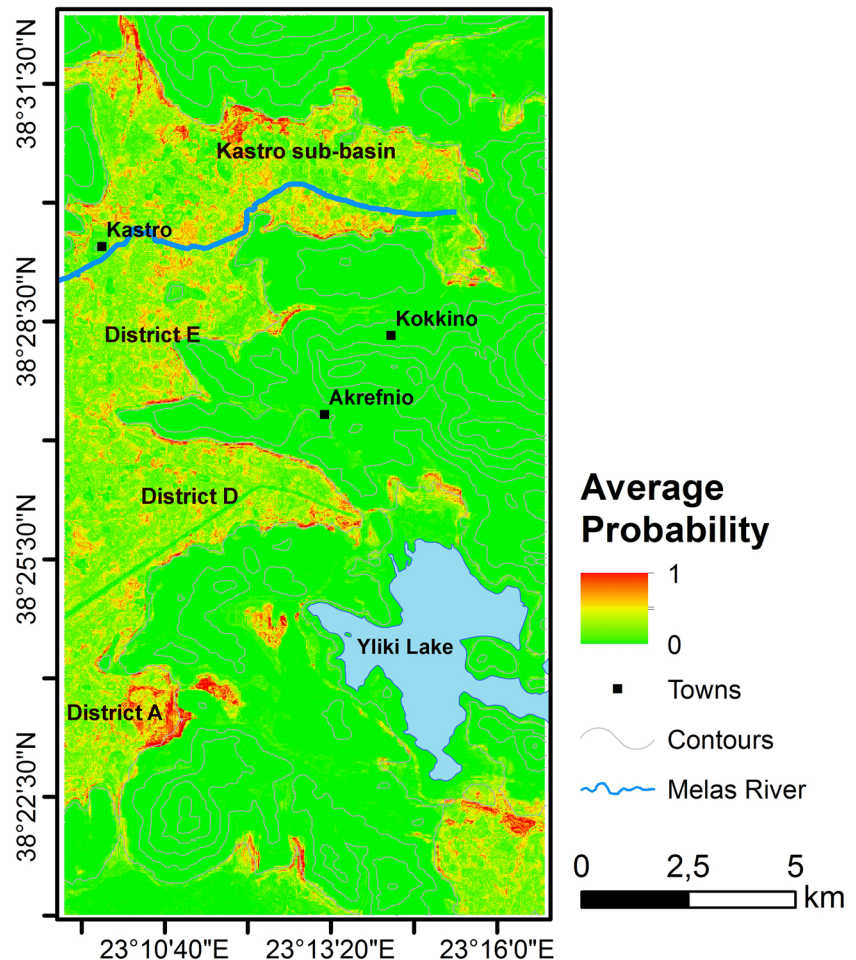
area. The vulnerability from statistical model mirrored the primary topographic properties of the Eastern Kopaida plain. This can be an important consideration for future studies as the borehole and analytical data density will be always limited to economic budgets while even high-resolution DEM are nowadays available at limited prices at a global scale.

The novelty of this contribution resided in the combination of two methods which historically have been kept apart between different scientific branches. The validation results generated from such different approaches proved once again that equally robust solutions of any complex physical process can be reached from numerous models. As such, bringing together the performances of PI and statistical methods may further improve the vulnerability assessment and better support future strategic schemes of groundwater resources management. A natural advancement could take place in future researches by keeping the well-established PI method as the reference model but increasing or decreasing the weighting factors following a spatial criterion derived by interpretation of statistical models.

References

Akpan AE, Ebong ED, Emeka CN (2015) Exploratory assessment of groundwater vulnerability to pollution in Abi, southeastern Nigeria, using geophysical and geological techniques. Env Mon Assess 187:156

Allen HD (1986) Late Quaternary of the Kopais Basin, Greece: sedimentary and environmental history. PhD Thesis, University of Cambridge, p 282

Almasri MN (2008) Assessment of intrinsic vulnerability to contamination for Gaza coastal aquifer, Palestine. J Environ Manag 88(4):577–593

Andreo B, Ravbar N, Vias JM (2009) Source vulnerability mapping in carbonate (karst) aquifers by extension of the COP method: application to pilot sites. Hydrogeol J 17:749–758

Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modeling. J Biogeogr 33(10):1677–1688

Beven KJ, Kirkby MJ (1979) A physically-based variable contributing area model of basin hydrology. Hydrol Sci Bull 24:43–69

Burkart MR, Kolpin DW, James DE (1999) Assessing groundwater vulnerability to agrichemical contamination in the Midwest US. Water Sci Technol 39(3):103–112

Chae G, Kim K, Yun S, Kim K, Kim S, Choi B, Kim H, Rhee CW (2004) Hydrogeochemistry of alluvial groundwaters in an agricultural area: an implication for groundwater contamination susceptibility. Chemosphere 55:369–378

Conoscenti C, Agnesi V, Angileri S, Cappadonia C, Rotigliano E, Märker M (2013) A GIS-based approach for gully erosion susceptibility modelling: a test in Sicily, Italy. Environ Earth Sci 70(3):1179–1195

Corani G, Mignatti A (2015) Robust Bayesian model averaging for the analysis of presence–absence data. Environ Ecol Stat 22:513–534

Davis AD, Long AJ, Wireman M (2002) KARSTIC: a sensitive method for carbonate aquifers in karst terrain. Environ Geol 42:65–72

Dörfliger N, Zwahlen F (1998) Practical guide: groundwater vulnerability mapping in karstic regions (EPIK). Environment in practice, Swiss Agency for the Environment, Forests and Landscape (SAEFL), Bern, p 56

European Water Directive (2000) Directive 2000/60/EC of the European parliament and of the council of 23 October 2000 establishing a framework for community action in the field of water policy. European Commission, Brussels

Giambelluca TW, Loague K, Green RE, Nullet MA (1996) Uncertainty in recharge estimation: impact on groundwater vulnerability assessments for the Pearl Harbor Basin, O'ahu, Hawai'i, USA. J Contam Hydrol 23:85–112

Goldscheider N (2005) Karst groundwater vulnerability mapping: application of a new method in the Swabian Alb, Germany. Hydrogeol J 13:555–564

Goldscheider N, Klute M, Sturm S, Hotzl H (2000) The PI method—a GIS based approach to mapping groundwater vulnerability with special consideration of karst aquifers. Z Angew Geol 46(3):157–166

Holting B, Haertle T, Hohberger K, Nachtigall K, Villinger E, Weinzierl W, Wrobel JP (1995) Konzept zur ermittlung der Schutzfunktion der Grundwasseruberdeckung. Geol Jb C63:5–24

Horn BKP (1981) Hillshading and the reflectance map. Proc IEEE 69(1):14–47

Jayasekera DL, Kaluarachchi JJ, Villholth KG (2011) Groundwater stress and vulnerability in rural coastal aquifers under competing demands: a case study from Sri Lanka. Env Mon Assess 176(1–4):13–30

Kaliraj S, Chandrasekar N, Simon Peter T, Selvakumar S, Magesh NS (2014) Mapping of coastal aquifer vulnerable zone in the south west coast of Kanyakumari, South India, using GIS-based DRASTIC model. Env Mon Assess 187:4073

Kavouri K, Plagnes V, Tremoulet J, Dörfliger N, Rejiba F, Marchet P (2011) PaPRIKa: a method for estimating karst resource and source vulnerability—application to the Ouysse karst system (southwest France). Hydrogeol J 19:339–353

Kumar P et al (2015) Index-based groundwater vulnerability mapping models using hydrogeological settings: a critical evaluation. Environ Impact Assess Rev 51:38–49

Lake IR, Lovett AA, Hiscock KM, Betson M, Foley A, Sünnenberg G, Evers S, Fletcher S (2003) Evaluating factors influencing groundwater vulnerability to nitrate pollution: developing the potential of GIS. J Environ Manag 68(3):315–328

Lombardo L, Cama M, Maerker M, Rotigliano E (2014) A test of transferability for landslides susceptibility models under extreme climatic events: application to the Messina 2009 disaster. Nat Hazards 74(3):1951–1989

Lombardo L, Cama M, Conoscenti C, Märker M, Rotigliano E (2015) Binary logistic regression versus stochastic gradient boosted decision trees in assessing landslide susceptibility for multiple-occurring landslide events: application to the 2009 storm event in Messina (Sicily, southern Italy). Nat Hazards 79(3):1621–1648

Mair A, El-Kadi AI (2013) Logistic regression modeling to assess groundwater vulnerability to contamination in Hawaii, USA. J Contam Hydrol 153:1–23

Malik P, Svasta J (1999) REKS: an alternative method of Karst groundwater vulnerability estimation. Hydrogeology and land use management, proceeding of the XXIX congress of the international association of hydrogeologists, Bratislava, Slovakia, pp 79–85

Margat J (1968): Vulnérabilité des nappes d'eau souterraine à la pollution—BRGM-Publication 68 SGL 198 HYD; Orléans

Masetti M, Poli S, Sterlacchini S (2007) The use of the weights-of-evidence modeling technique to estimate the vulnerability of groundwater to nitrate contamination. Nat Resour Res 16(2):109–119

Nohegar A, Riahi F (2014) The comparison of fuzzy drastic model and conventional drastic model to determine the most appropriate indicator of ground water vulnerability (Case Study: Sarkhoon Plain Aquifer). J Environ Stud 40(3):31–34

Nolan BT (2002) Relating nitrogen sources and aquifer susceptibility to nitrate in shallow ground waters of the United States. Ground Water 39(2):290–299

Pacheco FAL, Pires LMGR, Santos RMB, Sanches Fernandes LF (2015) Factor weighting in DRASTIC modeling. Sci Total Environ 505:474–486

Pagounis M, Gkertzos T, Gkatzogiannis A (1994) Hydrogeological research of Viotikos Kifissos River basin. Technical report, Institute of Geological and Mineral Research (IGME), Greece, p 94 **(in Greek)**

Parolo G, Rossi G, Ferrarini A (2008) Toward improved species niche modelling: Arnica montana in the Alps as a case study. J Appl Ecol 45(5):1410–1418

Pételet-Giraud E, Dörfliger N, Crochet P (2000) RISKE: methodology for multicriteria evaluation of the vulnerability of karst aquifers. Application to systems Fontanilles and Cent-Fonts Fontanilles (Herault, southern France). Hydrogéologie 4:71–88

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model 190(3–4):231–259

Plagnes V, Théry S, Fontaine L, Bakalowicz M, Dörfliger N (2005) Karst vulnerability mapping: improvement of the RISKE method. In: Proceedings of KARST 2005 Water Resources Environmental Problems Karst, vol 2, Belgrade-Kotor, Serbia, pp 123–140, 14–19 September 2005

Prasad RK, Singh VS, Krishnamacharyulu SKG, Banerjee P (2011) Application of drastic model and GIS: for assessing vulnerability in hard rock granitic aquifer. Env Mon Assess 176(1–4):143–155

Ravbar N, Goldscheider N (2007) Proposed methodology of vulnerability and contamination risk mapping for the protection of karst aquifers in Slovenia. Acta Carsol 36(3):461

Renner IW, Elith J, Baddeley A, Fithian W, Hastie T, Phillips SJ, Popovic G, Warton DI (2015) Point process models for presence-only analysis. Methods Ecol Evol 6(4):366–379

Sahragard HP, Chahouki MAZ (2015) An evaluation of predictive habitat models performance of plant species in Hoze soltan rangelands of Qom province. Ecol Model 309–310:64–71

Shaffer MJ, Lasnik K, Ou X, Flynn R (2001) NLEAP internet tools for estimating NO3-N leaching and N2O emissions. In: Shaffer MJ, Ma L, Hansen S (eds) Modeling carbon and N dynamics for soil management. Lewis Publishers, Boca Raton, pp 403–426

Soutter M, Musy A (1998) Coupling 1D Monte-Carlo simulations and geostatistics to assess groundwater vulnerability to pesticide contamination on a regional scale. J Contam Hydrol 32:25–39

Tagil S, Jenness J (2008) GIS-based automated landform classification and topographic, landcover and geologic attributes of landforms around the Yazoren Polje, Turkey. J Appl Sci 8:910–921

Tziritis E (2008) Hydrogeochemical—environmental study of east Kopaida—Yliki karstic system and assessment of vulnerability with the use of Geoinformatics. PhD thesis, University of Athens, Faculty of Geology and Geoenvironment

Tziritis E (2009) Groundwater and soil geochemistry of Eastern Kopaida region, (Beotia, central Greece). Cent Eur J Geosci 1(2):219–226

Tziritis E (2010) Assessment of NO3− contamination in a karstic aquifer, with the use of geochemical data and spatial analysis. Environ Earth Sci 60(7):1381–1390

Vias JM, Andreo B, Peles MJ, Carrasco F, Vadillo I, Jimenez P (2002) Preliminary proposal of a method for vulnerability mapping in carbonate aquifers. In: Carrasco F, Duran JJ, Andreo B (eds) Karst and environment. 2nd Nerja cave geological symposium, Nerja, Spain, 2002, pp 75–83

Vrba J, Zaporozec A (1994) Guidebook on mapping groundwater vulnerability. International contributions to hydrogeology (IAH), Hannover

Worrall F, Besien T, Kolpin DW (2002) Groundwater vulnerability: interactions of chemical and site properties. Sci Total Environ 299(1–3):131–143

Zwahlen F (2003) Vulnerability and risk mapping for the protection of carbonate (karst) aquifers, final report (COST action 620). European Commission, Directorate-General XII Science, Research and Development, Brussels, p 297