1 Table S1.Widely used Overlay and Index-based modelling techniques

	Method	Source	Equation	Details of Weight and Rating Systems		
1	DRASTIC	Aller et al., 1987	$DRASTIC_{Index} = \sum_{i=1}^{7} WiRi$ W_i and R_i=weight and rating for the i^{th} parameter, respectively.	**Feature**		**Weight**
				D	Depth to water	5
				R	Recharge	4
				A	Aquifer media	3
				S	Soil media	2
				T	Topography slope	1
				I	Impact of vadose zone	5
				C	Conductivity of the aquifer	3
2	GOD	Foster, 1987	$GOD_{Index} = GR \times OR \times DR$ G=groundwater occurrence including recharge, O=overlying lithology, and D=depth to groundwater of the method explained ahead, and subscript R=rating of the parameters	**GOD Parameter**		**Rating**
				(G)W confinement	Non aquifer	0
					Artesian	0.1
					Confined	0.2
					Semiconfined	0.3
					Unconfined with cover	0.4 - 0.6
					Unconfined	0.7 - 1
				(O)verlaying strata	Residual soil	0.4
					Limon alluvial; loess; shale, fine limestone	0.5
					Aeolian sand; siltite; tuff; igneous rock	0.6
					Sand and gravel; sandstone	0.7
					Gravel	0.8
					Limestone	0.9
					Fractured or karstic limestone	1
				(D)epth of GW (in m)	<2	1
					2-5	0.9
					5-10	0.8
					10-20	0.7
					20-50	0.6

<table>
<tr><td></td><td></td><td></td><td></td><td></td><td>50-100</td><td></td><td>0.5</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>>100</td><td></td><td>0.4</td></tr>
<tr><td rowspan="6">3</td><td rowspan="6">AVI</td><td rowspan="6">Van Stempvoort et al., 1993</td><td rowspan="6">$c = \sum_{i=1}^{n} \frac{d_i}{k_i}$
c = the hydraulic resistance given by AVI rating system; n = number of sedimentary units above the aquifer; d = thickness of each sedimentary unit, and k = estimated hydraulic conductivity of each sedimentary unit.</td><td colspan="2">Hydraulic resistance (year)</td><td>Log (c)</td><td>Vulnerability (AVI)</td></tr>
<tr><td colspan="2">0–10</td><td><1</td><td>Extremely high</td></tr>
<tr><td colspan="2">10–100</td><td>1-2</td><td>High</td></tr>
<tr><td colspan="2">100–1,000</td><td>2-3</td><td>Moderate</td></tr>
<tr><td colspan="2">1,000–10,000</td><td>3-4</td><td>Low</td></tr>
<tr><td colspan="2">>10,000</td><td>>4</td><td>Extremely low</td></tr>
<tr><td rowspan="7">4</td><td rowspan="7">ISIS</td><td rowspan="7">Civita & Regibus, 1995</td><td rowspan="7">$I_v = (R_{inf} \times Cinf) + (R_{su} \times Csus \times Csu) + (R_{ins} \times Csi \times Cins) + (R_{sat} \times Css \times Csat)$
R_{inf} = rating values for ranges of the net recharge; C_{inf} = infiltration coefficient dependent on land use; R_{su} = rating values for the soil media; C_{sus} = soil coefficient dependent on land use; C_{su} = weighting coefficient dependent on soil thickness; R_{ins} = rating values assigned to the vadose zone; C_{si} = weighting coefficient dependent on the unsaturated-zone lithology and thickness; C_{ins} = vadose zone coefficient dependent on land use; R_{sat} = rating values assigned to aquifer media; C_{ss} = weighting coefficient dependent on the aquifer thickness; and C_{sat} = aquifer coefficient dependent on land use.</td><td colspan="3">Range</td><td>Index</td></tr>
<tr><td colspan="3">141–180</td><td>Extreme</td></tr>
<tr><td colspan="3">124–140</td><td>Very high</td></tr>
<tr><td colspan="3">88–123</td><td>High</td></tr>
<tr><td colspan="3">64–87</td><td>Medium</td></tr>
<tr><td colspan="3">44–63</td><td>Low</td></tr>
<tr><td colspan="3">24–43</td><td>Very low</td></tr>
</table>

<table>
<tr><td rowspan="8">5</td><td rowspan="8">SINTACS</td><td rowspan="8">Civita & Maio, 1997</td><td rowspan="8">$I_v = \sum_{i=1}^{7}\sum_{i=1}^{n}(Pi \times Wj)$

P_i = rating of the ith parameter, and W_j = associated weight of the jth weight classification.</td><td colspan="2">Parameter</td><td>Weight</td></tr>
<tr><td>S</td><td>Soggiacenza (depth to water)</td><td>5</td></tr>
<tr><td>I</td><td>Infiltracione efficace (net recharge)</td><td>4</td></tr>
<tr><td>N</td><td>Non-saturo–efecto di autodepurazionedel- (impact of vadose zone)</td><td>5</td></tr>
<tr><td>T</td><td>Tipología della copertura (type of soil cover)</td><td>4</td></tr>
<tr><td>A</td><td>Acquifero–caratteristiche idrogeologichedel- (lithology of the aquifer)</td><td>3</td></tr>
<tr><td>C</td><td>Conducitibilitá hidráulica dell’acquifero (hydraulic conductivity of the aquifer)</td><td>3</td></tr>
<tr><td>S</td><td>Superficie topográfica–acclivitá della- (slope)</td><td>2</td></tr>
<tr><td rowspan="5">6</td><td rowspan="5">EPIK</td><td rowspan="5">Doerfliger et al., 1999</td><td rowspan="5">$Fp_i = aEi + bPi + cI_i + dKi$

Fp_i = protection factor pertaining to the ith cell; E_i, P_i, I_i, K_i = weights considered for the ith cell (range of parameters depending upon their impact on the pollution potential are given in Doerfliger et al., 1999); a, b, c, d = attribute relative weights having values of 3, 1, 3 and 2, respectively; and i = 1, .., n (grid cell number).</td><td colspan="2">Parameter</td><td>Weight</td></tr>
<tr><td>E</td><td>Epikarst (α)</td><td>3</td></tr>
<tr><td>P</td><td>Protective cover (β)</td><td>1</td></tr>
<tr><td>I</td><td>Infiltration conditions (γ)</td><td>3</td></tr>
<tr><td>K</td><td>Karst network development (δ)</td><td>2</td></tr>
<tr><td rowspan="7">7</td><td rowspan="7">PI</td><td rowspan="7">Goldscheider et al., 2000</td><td rowspan="7">$p = P \times I$

p = protection factor, P = protective cover, and I = infiltration conditions.</td><td colspan="2">P-map
Effectiveness of Protective Cover</td><td colspan="2">I-map
Degree of Bypassing</td><td colspan="2">Vulnerability Map
Vulnerability of uppermost Aquifer</td></tr>
<tr><td>P-factor</td><td>verbal description</td><td>I-factor</td><td>verbal description</td><td>π-factor</td><td>verbal description</td></tr>
<tr><td>1</td><td>very high</td><td>0-0.2</td><td>very high</td><td>0-1</td><td>extreme</td></tr>
<tr><td>2</td><td>high</td><td>0.4</td><td>high</td><td>>1-2</td><td>high</td></tr>
<tr><td>3</td><td>moderate</td><td>0.6</td><td>moderate</td><td>>2-3</td><td>moderate</td></tr>
<tr><td>4</td><td>low</td><td>0.8</td><td>low</td><td>>3-4</td><td>low</td></tr>
<tr><td>5</td><td>very low</td><td>1.0</td><td>very low</td><td>>4-5</td><td>very low</td></tr>
</table>

<table>
<tr><td>8</td><td>GALDIT</td><td>Chachadi & Ferreira, 2005</td><td>$GALDIT\ Index\ =\ \frac{\sum_{i=1}^{6}(W_i \times R_i)}{\sum_{i=1}^{6} W_i}$

W_i = weight of the i^{th} indicator, and R_i=rating of the i^{th} indicator.</td><td colspan="6"><table>
<tr><th colspan="2">Parameter</th><th>Weight</th></tr>
<tr><td>G</td><td>Groundwater Occurrence (Aquifer Type)</td><td>1</td></tr>
<tr><td>A</td><td>Aquifer Hydraulic Conductivity</td><td>3</td></tr>
<tr><td>L</td><td>Height of Groundwater Level above Sea Level</td><td>4</td></tr>
<tr><td>D</td><td>Distance from the Shore</td><td>4</td></tr>
<tr><td>I</td><td>Impact of existing status of Seawater Intrusion</td><td>1</td></tr>
<tr><td>T</td><td>Thickness of Aquifer being Mapped</td><td>2</td></tr>
</table></td></tr>
<tr><td>9</td><td>COP</td><td>Vías et al., 2006</td><td>$COP_{index} = C \times O \times P$

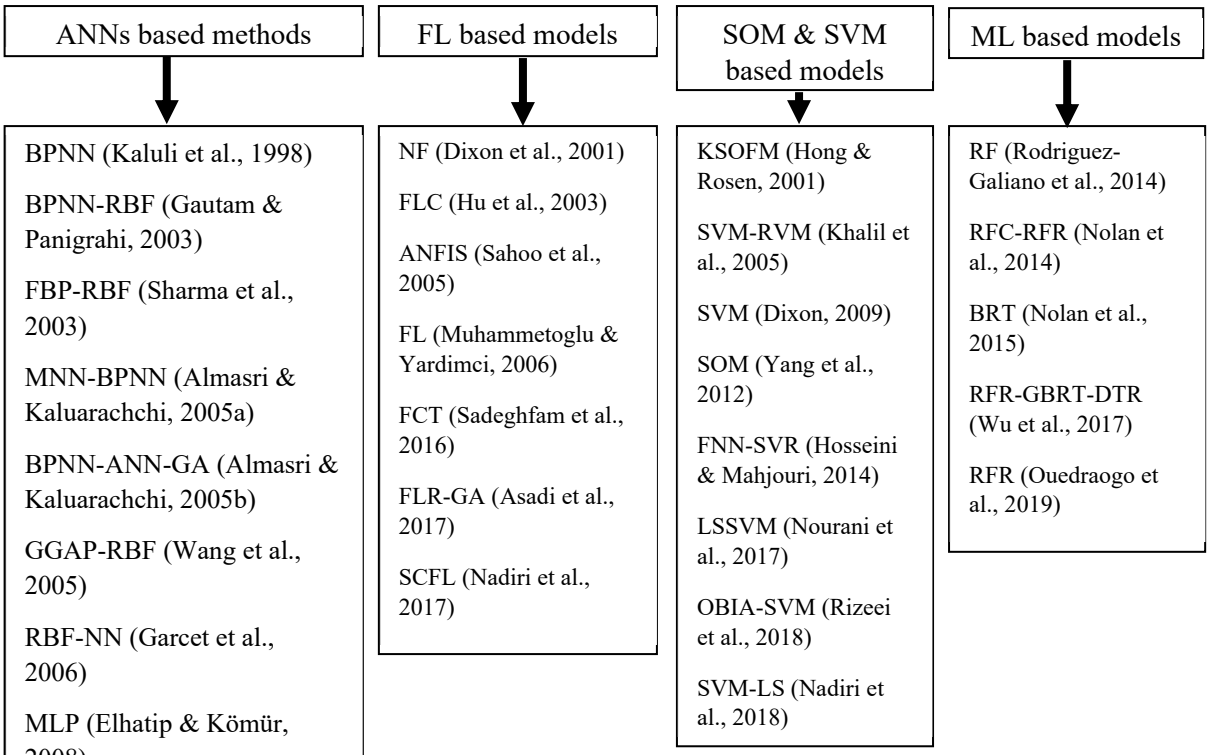
COP_{index} = intrinsic vulnerability of the karst aquifer, C = concentration, O = overlying layers, and P = precipitation.</td><td colspan="6"><table>
<tr><th colspan="2">O Factor (Overlaying layers)</th><th colspan="2">C Factor (Concentration of flow)</th><th colspan="2">P Factor (Precipitation)</th><th rowspan="2">COP Index</th><th rowspan="2">Vulnerability classes</th></tr>
<tr><th>O Score</th><th>Protection value</th><th>C Score</th><th>Reduction of protection</th><th>P Score</th><th>Reduction of protection</th></tr>
<tr><td>1</td><td>Very low</td><td>0-0.2</td><td>Very high</td><td>0.4-0.5</td><td>Very high</td><td>0-0.5</td><td>Very high</td></tr>
<tr><td>2</td><td>Low</td><td>0.2-0.4</td><td>High</td><td>0.6</td><td>High</td><td>0.5-1</td><td>High</td></tr>
<tr><td>2-4</td><td>Moderate</td><td>0.4-0.6</td><td>Moderate</td><td>0.7</td><td>Moderate</td><td>1-2</td><td>Moderate</td></tr>
<tr><td>4-8</td><td>High</td><td>0.6-0.8</td><td>Low</td><td>0.8</td><td>Low</td><td>2-4</td><td>Low</td></tr>
<tr><td>8-15</td><td>Very high</td><td>0.8-1.0</td><td>Very low</td><td>0.9-1.0</td><td>Very low</td><td>4-15</td><td>Very low</td></tr>
</table></td></tr>
<tr><td>10</td><td>PaPRIKa</td><td>Dorfliger & Plagnes, 2009</td><td>$V_g = (i \times I) + (r \times R) + (p \times P) + (k \times Ka)$

V_g=resource vulnerability map, P=protection factor, R=reservoir factor, I=infiltration factor, and Ka=karstification factor, and i, r, p and k are affecting weights with their sum equal to 1.</td><td colspan="6"><table>
<tr><th colspan="2" rowspan="2">Factor</th><th colspan="2">Rating (%)</th></tr>
<tr><th>For karst catchment</th><th>For non karst catchment</th></tr>
<tr><td>P</td><td>Protection (Overlaying-layer, Epikarst, soil cover, unsaturated zone)</td><td>20–25</td><td>50</td></tr>
<tr><td>R</td><td>Rock type</td><td>15–20</td><td>-</td></tr>
<tr><td>I</td><td>Infiltration</td><td>40-30</td><td>50</td></tr>
<tr><td>Ka</td><td>karstification degree</td><td>20–30</td><td>-</td></tr>
</table></td></tr>
<tr><td>11</td><td>CORE</td><td>Pavlis & Cummins, 2014</td><td>$CORE_{Score} = O \times C \times R \times E$

O = Overlaying layers; C = Concentrated flow; R = Recharge; E = Extreme events</td><td colspan="6"><table>
<tr><th colspan="2">Concentration of flow (C factor)</th><th colspan="2">Overlaying layers (O factor)</th><th colspan="2">Precipitation (Rainfall & Extreme events/ RE factor)</th><th rowspan="2">CORE vulnerability index</th></tr>
<tr><th>C Score</th><th>Reduction of protection</th><th>O Score</th><th>Protection</th><th>RE Score</th><th>Reduction of protection</th></tr>
<tr><td>0-0.2</td><td>Very High</td><td>1</td><td>Very Low</td><td>0.4-0.5</td><td>Very High</td><td>Very High</td></tr>
</table></td></tr>
</table>

				0.2-0.4	High	2	Low	0.5-0.6	High	High
				0.4-0.6	Moderate	3-5	Moderate	0.6-.07	Moderate	Moderate
				0.6-0.8	Low	5-7	High	0.7-0.8	Low	Low
				0.8-1.0	Very Low	8	Very High	>0.8	Very Low	Very Low

2



ANNs: Artificial Neural Networks; **BPNN:** Back Propagation Neural Networks; **RBF:** Radial Basis Function; **FBP:** Fast Back Propagation; **GA:** Genetic Algorithm; **MNN:** Modular Neural Network; **GGAP**: Generalized Growing and Pruning; **RBF-NN**: Radial Basis Function Neural Network; **MLP**: Multi-Layer Perceptron; **LM**: Levenberg – Marquardt; **BP:** Back Propagation; **FFBP**: Feed Forward Back Propagation; **GRNN**: Generalized Regression Neural Network; **FFNN**: Feed Forwarded Neural Network; Bayesian NN: Bayesian Neural Network; **SCMAI**: Supervised Committee Machine with Artificial Intelligence; **BP-BA:** Back Propagation - Bee Algorithm; **PC:** Principle Component.

NF: Neuro-Fuzzy; **FLC:** Fuzzy Logic Control system; **ANFIS:** Adaptive-Neuro Fuzzy Inference System; **FL:** Fuzzy Logic; **FCT:** Fuzzy Catastrophe Theory; **FLR-GA:** Fuzzy Logic Regression-Genetic Algorithm; **SCFL:** Supervised Committee model combined with FL.

SOM: Self-Organized Map; **KSOFM**: Kohonen Self-Organizing Feature Maps; **SVM-RVM:** Support Vector Machine - Relevance Vector Machines; **FNN-SVR:** Fuzzy Neural Network – Support Vector Regression; **LSSVM:** Least Square Support Vector Machine; **OBI:** Object-Based Image Analysis; **LS:** Least Square.

ML: Machine Learning; **RF:** Random Forest; **RFC:** Random Forest Classification; **RFR:** Random Forest Regression; **BRT:** Boosted Regression Tree; **GBRT:** Gradient Boosted Tree Regression; **DTR**: Decision Tree Regression

Figure 1. Statistical/Soft computing/Machine learning methods for GW contamination

4

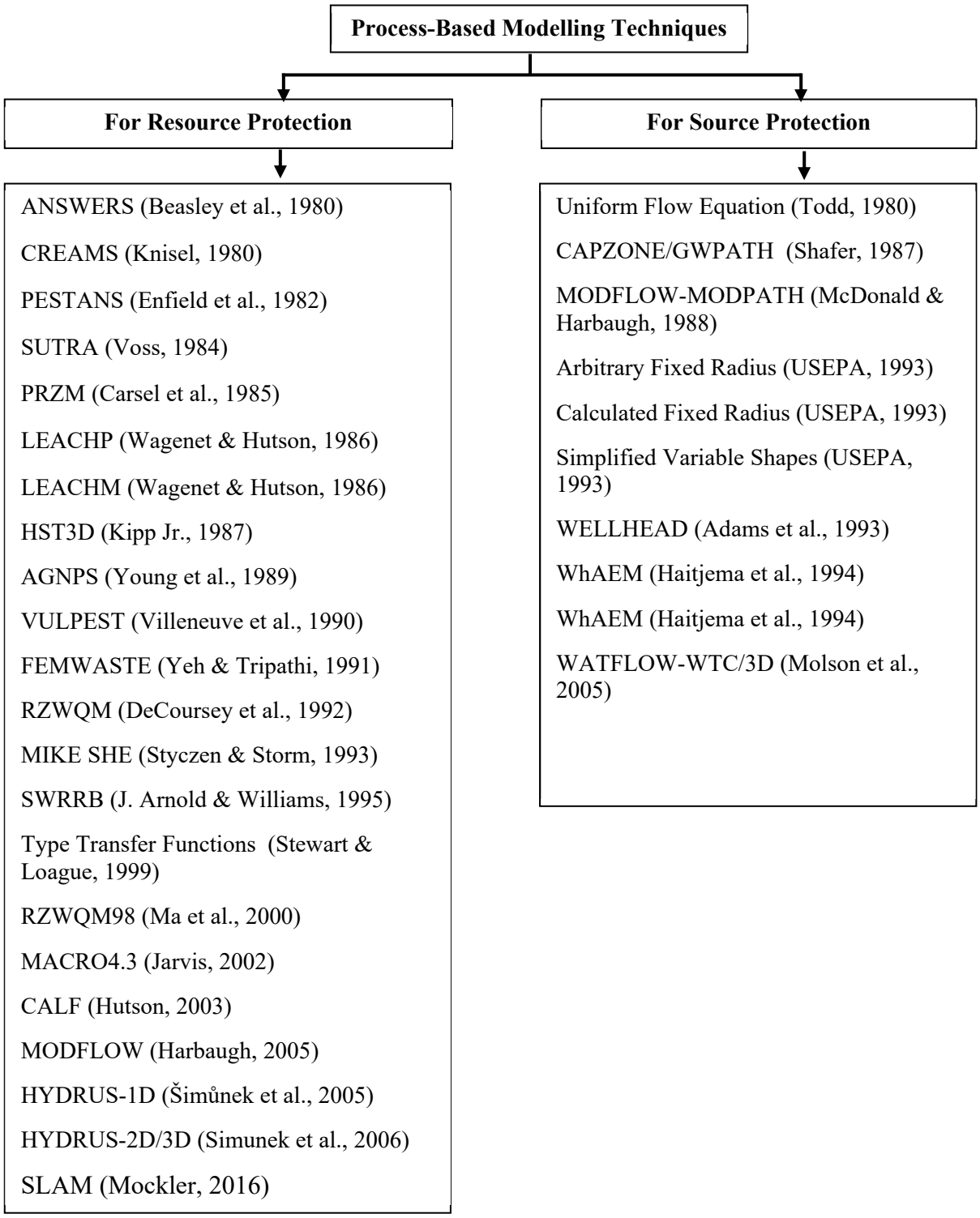


Figure 2.Widely used Process-Based modelling techniques for GW contamination assessment

5