Chapter 14
Case Study: Watershed Analysis (Hydrology)

14.1 Overview

Twenty-one watersheds were characterized and ranked on the basis of two multiple indicator systems, level 1 (least expensive) and level 2 (expensive). Furthermore, level 3 indicators are defined which need investigations in the field and are pretty expensive and only six watersheds are characterized.

Composite indicators are defined on the basis of level 1 indicators, called LSI, and level 2 indicators, called SWR. LSI and SWR are thought of as two different means to rank the watersheds with respect to the environmental health. The indicators on the three levels are considered as proxies to describe the abstract and not measurable concept of "environmental health."

One task in the poset analysis is the verification or the falsification of the two composite indicators.

Whereas LSI is fairly justified by the partial order analysis, the SWR needs more attention, as its weighting scheme differs from that derived from a partial order analysis, applying the POSAC method (Chapter 3). Furthermore, it is observed that the partial order based on level 1 indicators differs remarkably from that based on level 2 indicators. Finally, it is striking that a proximity analysis (see Chapter 10) favors level 1 indicators as proxies of level 3, albeit for this study only six watersheds were available.

By the sensitivity study, it turns out that the indicator "impervious surface in watersheds" was most important under the level 1 indicators, whereas "invasive cover class" was most important under level 2 indicators.

As in the analysis of child development (Chapter 12), separated subsets are found, which are an outcome of a partial order analysis of the indicators of level 1. The indicator IMP (impervious surface in watersheds), CORFOR (percentage of total forest that is core forest in watershed) as well as FOR (percentage of forest in watersheds) explain this separation.

Special attention was given to the level 3 (environmental chemistry and biology) indicators: An examination of NO_3 vs biological parameters renders strongly differing sensitivity to the weights (of NO_3 vs biological parameters): The watershed

R. Brüggemann, G.P. Patil, *Ranking and Prioritization for Multi-indicator Systems*, Environmental and Ecological Statistics 5, DOI 10.1007/978-1-4419-8477-7_14,

"Saint Mary's A" (SM) is almost insensitive, whereas the ranking of "back river" (BR) varies strongly with the weights.

14.2 Background and Data Matrix

The data set of 21 watersheds obtained from the Atlantic Slope Consortium (ASC) will be analyzed with the goal of determining an accurate ranking of the health of the watersheds. See Brooks et al. (2007). As in Chapter 13 (bridge study), the data set has different levels of indicators, grouped into level 1 to level 3, increasing in the quality and accuracy of the data as well as in the amount of cost and effort needed to obtain the data:

Level 1: Landscape assessment using satellite data is the easiest to access and the least expensive (data matrix, see Table A.14).

Level 2: Rapid field assessment is obtained from on-site sampling. Certain level of expertise is involved in the field assessment (data matrix, see Table A.15). Generally, level 2 data is relatively inexpensive compared to the level 3 data.

Level 3: Intensive field assessment needs to be purchased from the US Environmental Protection Agency (EPA). It is the most expensive and best quality of data among the three levels. Due to the money and effort in the procedure of obtaining this data, it is available for only six watersheds (data matrix, see Table A.16).

The orientation of all these indicators is as follows: The higher the indicator value, the better the watershed.

We provide an overview of the three-level indicator system in Table 14.1 and the names of the watersheds in Table 14.2.

The investigators combined the five level 1 indicators and the seven level 2 indicators to indices: The composite level 1 index is called the landscape index (LSI):

$$\mathrm{LSI} = [\mathrm{FOR} + (\mathrm{IMP} + \mathrm{LDI})/2 + (\mathrm{MPAT} + \mathrm{CORFOR})/2]/3 \tag{14.1}$$

The composite level 2 index is called stream–wetland–riparian (SWR) index:

$$\mathrm{SWR} = (1/4)^{*}(\mathrm{FW} + \mathrm{SHA} + \mathrm{IR} + \mathrm{SS}) \tag{14.2}$$

$$\mathrm{FW} = (1/4)^{*}(\mathrm{BUW} + \mathrm{BA} + \mathrm{INV} + \mathrm{FPWL}) \tag{14.3}$$

The values of both indices can be found in the appendix (Table A.17).

Table 14.1 Meaning of the indicators of the three levels

Level 1		Level 2		Level 3	
Indicator	Definition	Indicator	Definition	Indicator	Definition
FOR	Percentage of forest in watershed	BUF	Buffer score	BIBI	Benthic IBI (index of biological integrity)
LDI	Landscape density index in watershed	IR	Incision ratio	FIBI	Fish IBI
IMP	Percentage of impervious surface in watershed	BA	Basal area of trees	NO_3	Concentration of nitrate[a]
MPAT	Mean forest patch size in watershed	INV	Invasive cover class		
CORFOR	Percentage of total forest that is core forest in watershed	SHA	Stream habitat assessment score		
		SS	Number of stream stressors		
		FPWL	Number of floodplain–wetland stressors		

[a]The orientation of NO_3 has to be reversed because a large value indicates a bad state of the watershed

Table 14.2 Watershed names and their identifier

Watershed	Identifier	Watershed	Identifier
Back River	BR	Conodoguinet A	CA
Cattail Creek	CC	Grindle Creek	GC
Gwynn Falls	GF	Little Contentnea	LC
Saint Mary's A	SM	Mantua	Ma
Southeast Creek	SC	Middle Creek	MC
Upper Patuxent	UP	Middle River	MiR
Ahoskie	Ah	Pamunkey	Pa
Buffalo Creek	BC	Repaupo	Re
Chickahominy	Ch	White Deer Creek	WDC
Christian Creek	ChC	Wisconisco	Wi
Clearfield Creek	ClC		

14.3 Partial Order Analysis, Based on Level 1

14.3.1 Hasse Diagram

X: the set of 21 watersheds
IB: the set of five indicators of level 1 ([0,1]-normalized data)
Orientation: The larger the values of the indicators, the better the environmental health.

Once again five indicators are used as proxies for the abstract principle "environmental health" and the Hasse diagram (Fig. 14.1) shows how the watersheds can be positioned with respect to these five indicators.

Figure 14.1 shows the following:

- There are six levels: Strolling up the Hasse diagram, we find watersheds of increasing better state with respect to the abstract principle of environmental health.
- The incomparabilities tell us that a certain state with respect to the environmental health is realized by profiles (Chapter 3) crisscrossing each other.
- There is a least element, BR, having values with respect to all five indicators which are less than those of any other watershed.
- Shape: Approximately rectangular, the disparity in the values of the five indicators does not strongly vary with the levels.

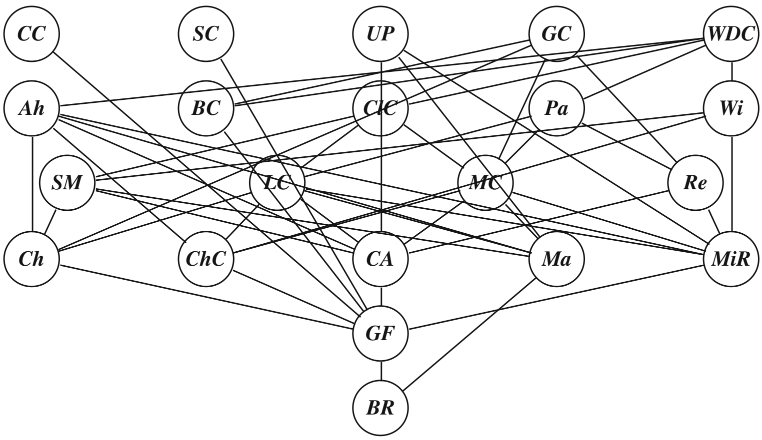

Fig. 14.1 Hasse diagram of (*X*, IB)

The Hasse diagram of Fig. 14.1 also supports some comparative decisions:

- A crude evaluation follows from the membership of watersheds to one of these order theoretical levels.
- There are some chains having at least four elements, allowing to order the watersheds uniquely without the use of LSI. For example, BR < GF < CA < LC < CC < GC, or BR < Ma < MC < CC < WDC. A list of all paths between two endpoints can be obtained by PyHasse.

The indicators of level 1 have different impact on the Hasse diagram as can be obtained from a sensitivity study (Chapter 4). For level 1 indicators, the following sequence (ordered for decreasing importance for a Hasse diagram) is found: IMP >> FOR > CORFOR > LDI ≅ MPAT.

CAM = 0.5. Therefore, any new indicator added or any deletion of an indicator may remarkably change the partial order (see Chapter 4).

14.3.2 Antagonism

Let X be the set of watersheds and $X_{\mathrm{res}} := \mathrm{X}-\{\mathrm{GF}, \mathrm{BR}\}$, then {CC, SC} and $X_{\mathrm{res}} = \{\mathrm{UP}, \mathrm{GC}, \mathrm{WDC}, \mathrm{Ah}, \mathrm{BC}, \mathrm{ClC}, \mathrm{Pa}, \mathrm{Wi}, \mathrm{SM}, \mathrm{LC}, \mathrm{MC}, \mathrm{Re}, \mathrm{Ch}, \mathrm{ChC}, \mathrm{CA}, \mathrm{Ma}, \mathrm{MiR}\} = X' \cup \{\mathrm{Ma}\}$ are separated subsets (Chapter 5). By the tools provided by WHASSE, we find that IMP and CORFOR explain the separation at 94.1%. The complete separation (100%) is obtained if indicator FOR is included. Hence AIB = {FOR, IMP, CORFOR}. The scatter plot based on the 94.1 approximation is shown in Fig. 14.2.

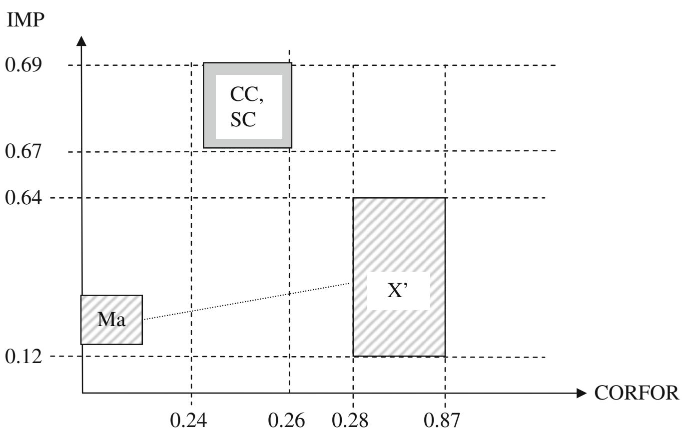

Fig. 14.2 Scatter plot explaining 94.1% antagonism between {SC,CC} and X_{res}. Watershed Ma belongs to X_{res} (indicated by a *broken line*)

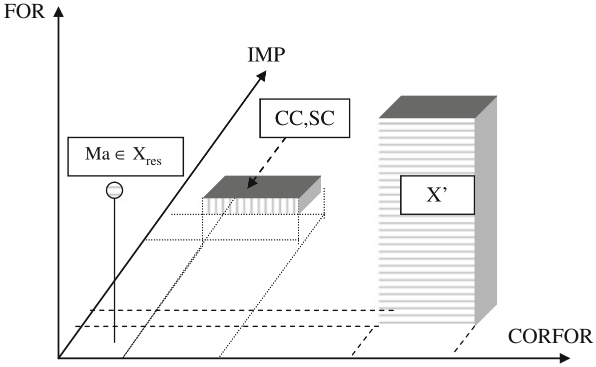

Fig. 14.3 3D schematic view on the results of the antagonism study

As one can see in Fig. 14.2, the separation {CC, SC} from X_{res} is not complete, because $\text{Ma} <_{\{\text{CORFOR, IMP}\}} \text{CC}$ and $\text{Ma} <_{\{\text{CORFOR, IMP}\}} \text{SC}$. The third indicator FOR does the job of a complete separation. FOR(CC) < FOR(Ma) as well as FOR(SC) < FOR(Ma). Figure 14.3 summarizes schematically the results of the antagonism study.

We see from Fig. 14.3 that the range of FOR(X_{res}-{Ma}) includes that of {CC, SC}. This is consistent with the finding of Section 5.5.3.

14.4 Partial Order Point of View, Level 2 Indicators

The Hasse diagram of (X, $\text{IB}_{\text{level2}}$)

Object set X: 21 watersheds as before
$\text{IB}_{\text{level2}}$: Seven indicators of level 2.
Orientation: The larger the value, the better the state of the watersheds with respect to environmental health.

The Hasse diagram can be inspected in Fig. 14.4.

One may expect that the higher number of level 2 indicators leads to more contradictions in the data. This is indeed the case: Instead of six levels in the case of level 1 indicators, there are now only three. Furthermore, we see the following:

- $\text{ISO} = \{\text{BC}, \text{ClC}\}$
- WDC, GF, and GC are articulation points. Deletion of any of the corresponding rows from the data matrix generates at least one more isolated element. Deletion of WDC from the data matrix would lead to four more components in the Hasse diagram.

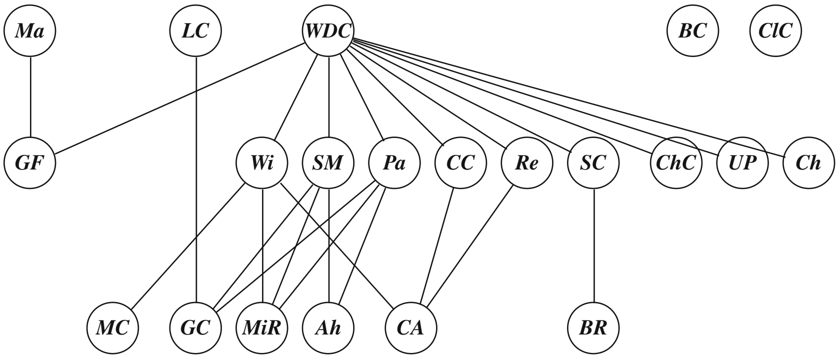

Fig. 14.4 Hasse diagram of (X, IB_{level2})

- MAX = {Ma, LC, WDC, BC, ClC}. The only maximal element which is also found through level 1 indicators is WDC.
- The proximity analysis between (X, IB_{level1}) and (X, IB_{level2}) (Chapter 10) renders fraction (isotone) = 0.08, fraction (indifferent) = 0.92, fractions of antitone, weak isotone, and equivalent = 0. We conclude that level 2 indicators do not contradict those of level 1. However, they strongly reveal different information about the watersheds.

14.4.1 Attribute-Related Sensitivity

Figure 14.5 shows the result of the sensitivity analysis.

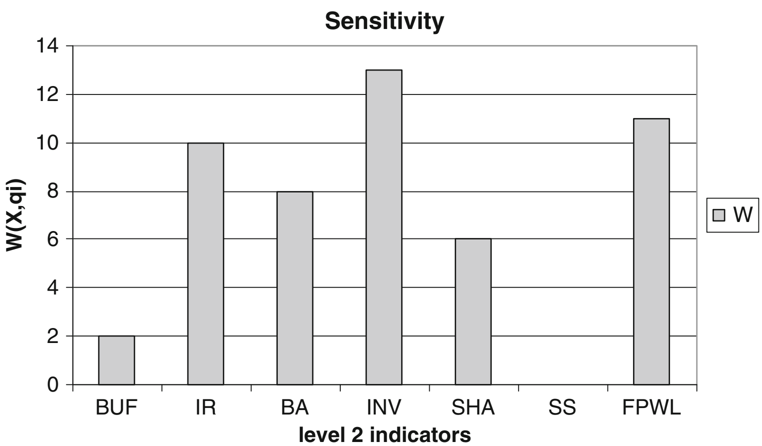

Fig. 14.5 Attribute-related sensitivity, level 2 indicators. The most important indicator is INV and the least one is SS

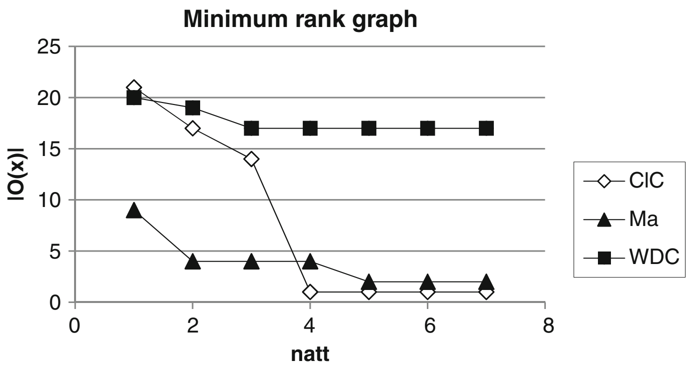

Fig. 14.6 Minimum rank graph of ClC, WDC and Ma in (X, IB_{Level2}) (see Chapter 4)

14.4.2 Minimum Rank Graph

We take ClC, WDC, and MA (which are maximal elements of (X, IB_{level2}) and determine their minimum rank graphs (Fig. 14.6).

In Fig. 14.6, we see the following:

- WDC has a slightly worse position than ClC but is pretty invariant, whereas ClC moves down through adding indicators.
- ClC has a steep gradient when the fourth indicator is added.
- The watershed Ma is in a middle position if only the fourth indicator, INV, is considered. All the other additional indicators affect Ma only slightly.

For a closer interpretation, we show the data profiles of ClC, WDC, and Ma in Fig. 14.7.

We see that each of the three watersheds has at least one indicator, where the watershed is better than the other two. In the case of indicator 4 of level 2, ClC is slightly better than WDC; however, the numerical difference is too small to be visualized.

WDC: Three indicators of WDC are 1 or nearly 1, inclusive of the most important indicator INV and the lowest value is 0.72. Hence adding the indicators has little influence on the position of WDC.

ClC: Adding the indicator BA to the data matrix must eliminate all successors of ClC, because BA(ClC) is near 0 (note: normalized data). All other watersheds have better values. Better values in the remaining indicators of ClC cannot increase the number of successors.

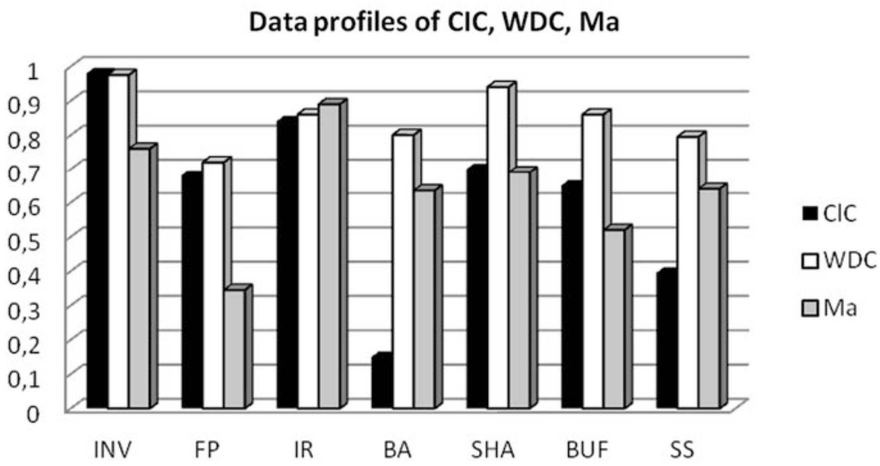

Fig. 14.7 Data profiles (normalized) of three maximal elements in $(X, \mathrm{IB}_{\mathrm{level2}})$, the level 2 indicators are ordered by $W(q_i)$

Ma: This watershed is a maximal element because of the good value of indicator IR. With respect to the indicator INV, the watershed Ma has only a medium to fair value.

Hence the minimum rank graph starts at rather low values. Adding the next important indicator FP excludes many successors, because Ma has here its lowest value. Therefore, a strong negative slope appears. As in the case of ClC, the better values in the remaining indicators cannot increase the number of successors.

14.5 Analysis Including the Level 3 Indicators

14.5.1 Hasse Diagrams

The Hasse diagram for the level 3 indicators is shown in Fig. 14.8.

Figure 14.8 shows the important role of the indicator NO_3 for the Hasse diagram.

14.5.2 Indicator NO_3 vs Biological Indicators FIBI and BIBI

Figure 14.8 motivates to study in more detail the role of NO_3 vs the biological indicators. Hereto we define

$$\mathrm{IBI} := 0.5^*(\mathrm{BIBI} + \mathrm{FIBI}) \text{ and } \varphi := g^*\mathrm{IBI} + (1-g)^*\mathrm{NO}_3 \tag{14.4}$$

We perform the stability analysis (Fig. 14.9).

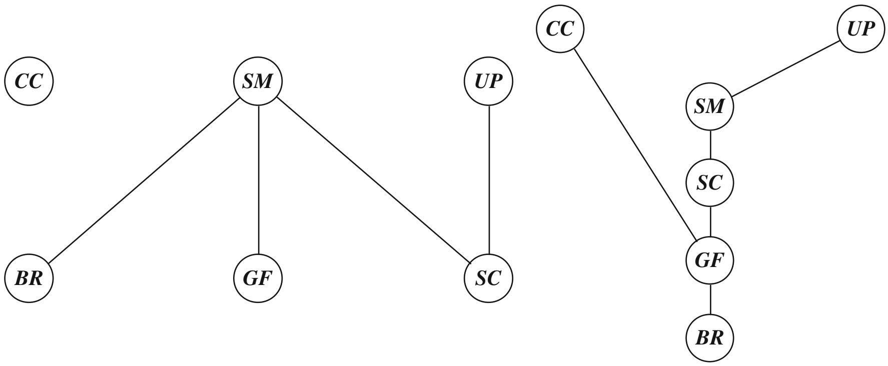

Fig. 14.8 (*LHS*) Hasse diagram for level 3 indicators; (*RHS*) FIBI and BIBI

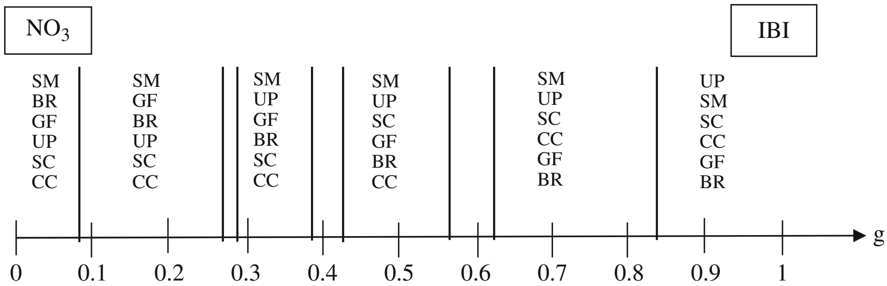

Fig. 14.9 Stability plot for the study of six watersheds and level 3 indicators and the composite indicator: $\varphi = g^*(0.5^*\mathrm{BIBI} + 0.5^*\mathrm{FIBI}) + (1-g)^*\mathrm{NO}_3$. *Vertical linear orders* given with enough space between two subsequent g_c values

In Fig. 14.9, the results of METEOR and stability analysis (see Chapter 7) are shown. We identify eight crucial weights. The stability fields together with the linear orders are displayed. With $g = 0$ (LHS), the linear order due to indicator NO_3 is obtained and with $g = 1$ (RHS), that of indicator IBI is obtained. One can see that

- the high height in the linear orders of watershed SM is rather stable; only if the weight g is larger than 0.84, it changes its position with watershed UP.
- the watershed CC has the lowest position until $g \approx 0.6$ but remains in the lower part of the ranking.
- the watershed SC has height = 2 for $0 \leq g \leq 0.39$. For all larger values of g, the watershed SC remains in the height position 4. The range of g values around 0.39 is a hot spot for SC.
- watershed GF is pretty sensitive to the amount how far NO_3 is mixed with IBI.

- there are two stability fields which are rather large, one between $g^* \approx 0.1$ and $g^* \approx 0.3$ and another between 0.6 and 0.85. If the indicator NO_3 is to be included, then the ranking of the watersheds does not actually depend on weights taken from these two ranges.

14.6 Proximity Analysis of Level 1, 2, and 3 Indicators on the Basis of Six Watersheds

We conclude the partial order analysis of the wetlands with the focus on comparing the three sets of indicators on the basis of six watersheds. What we want to know is, which set of indicators is a better proxy for level 3 indicators when a comparative analysis is our focus. Therefore we analyze the following:

- level 3 vs level 2
- level 3 vs level 1
- level 1 vs level 2

We perform for all indicators a discretization ($K = 3$, $\min_i$ and $\max_i$ values taken from the data matrix). The resulting three Hasse diagrams were compared by means of the proximity analysis (Chapter 10). The results are shown in Fig. 14.10.

We see the following:

1. There is no contribution to antitone
2. There are no weak isotone contributions
3. Dominating is the frequency of indifferent matchings
4. Focusing on how well level 3 indicators are modeled by level 2 or level 1 indicators, the degree of isotone between level 3 and level 1 is 10, whereas that between level 3 and level 2 is only 2.

We conclude the following:

1. The indicators of the three levels do not contradict each other.
2. The indicators are sharp enough so that there do not appear combinations like $(\cong, <)$ (see Section 10.6). Therefore, the number of weak isotone equals 0.
3. Any indicator scheme has its own scientific value. Therefore, the objects are likely to be incomparable. Hence the contribution of "indifferent" is pretty high.
4. Taking into account that the results are based on only six watersheds and on a specific discretization scheme, we hypothesize that level 1 indicators are better suitable to be used as proxies for level 3 indicators compared to level 2 indicators.

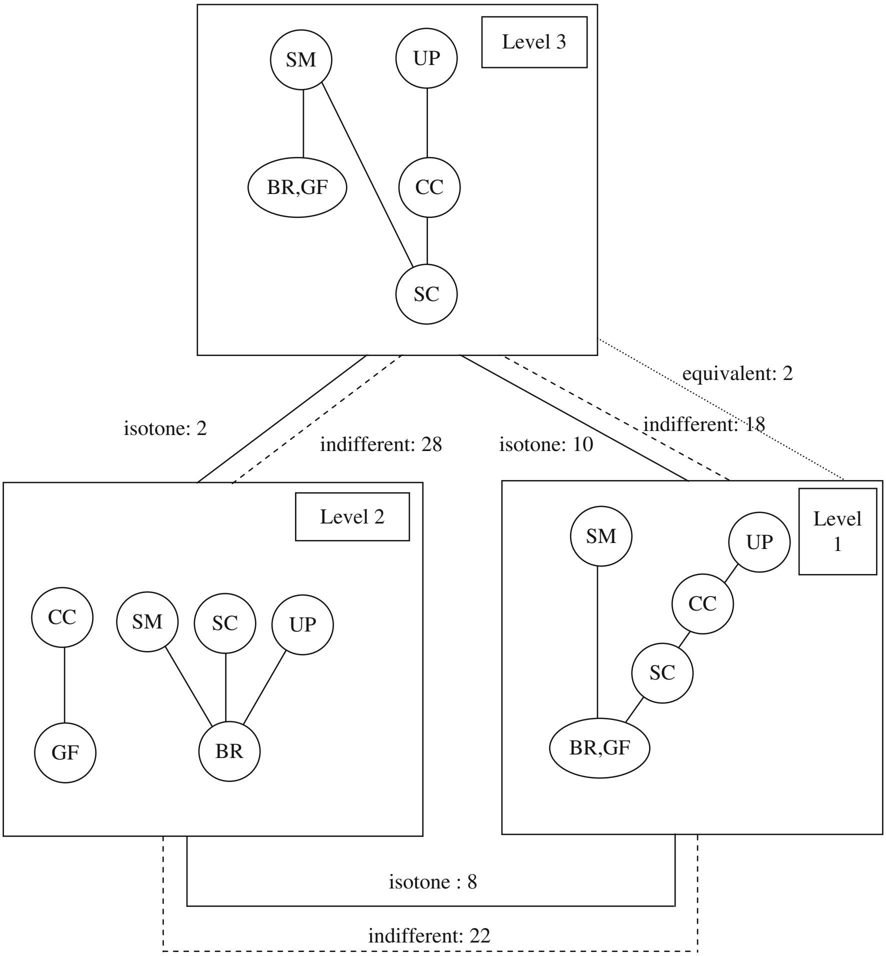

Fig. 14.10 Proximity analysis of the partial orders obtained from the three-level indicator systems for six watersheds

14.7 Analysis of LSI and SWR

14.7.1 Where Are We?

We have studied the Hasse diagrams and got an impression about the positions of the watersheds depending on the set of indicators without crunching the 5, 7, or 3 indicators into a composite indicator. All the three Hasse diagrams (level 1, level 2, and level 3) allow some comparison of the watersheds. Furthermore, we analyzed the indicator NO_3 vs. IBI as if we want to construct a composite indicator based on NO_3 and IBI, without having an idea about the weights. The order theoretical answer is to construct stability fields and hot spots in order to identify ranges of the weights, on the one hand, where some freedom in selecting the numerical value is and, on the other hand, where a slight variation will change the linear order of the watersheds.

However, the question is: Can we provide an alternative to the linear or the weak order a composite indicator provides? Here we do not want to be repetitive with apply methods explained in Chapter 10. Instead we will examine what the partial order tool POSAC (Chapter 3) has to offer to us.

14.7.2 Analysis with POSAC

As example, we take a closer look at the index SWR: We compare the weights given by the experts with weights which we develop from the data matrix itself. The tools we are applying is POSAC (see Chapter 3) and the concordance method to find the loadings, i.e., how far the original indicators contribute to the latent order variables of POSAC.

In this section we are closely following Patil (2001).

14.7.2.1 Loadings

As shown in Section 3.5, POSAC finds two latent order variables, LOV1 and LOV2 in short, and each object has a LOV1 and a LOV2 value corresponding to the POSAC diagram. We are interested in understanding the strength of the influence of the original indicators on the LOVs, the "loading." The loadings are computed for each indicator, and a loading gives a measure of similarity between the LOV and the data from a particular indicator. To allow for small deviations in the POSAC algorithm, we discretize both the LOVs and the original data into eight equidistant intervals. We compute a concordance value for each indicator and each of the both LOVs.

For the level 2 data set, 84.6% of the comparabilities are preserved by the two-dimensional POSAC model, and the two-dimensional POSAC diagram is in Fig. 14.11.

Table 14.3 shows the loadings concerning the level 2 indicators.

For level 2 the latent order variable LOV_1 is most impacted by indicators IR, which is the incision ratio, IR, followed by INV, which is invasive cover class, and SHA, the stream habitat assessment score. The latent order variable LOV_2 is most impacted by BUF, which is the buffer score of the watershed, and to a smaller extent by INV, SHA, and FPWL.

14.7.2.2 Derivation of Weights from the Data Matrix

Using the results of the analysis above on the loadings, we generate weights for the indicators which are solely based on the data matrix. Let I_i be one of the indicators and a_{ij} the loading for the jth LOV:

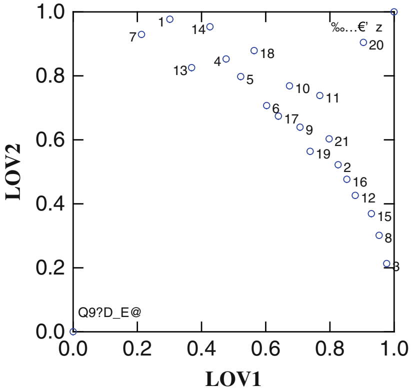

Fig. 14.11 POSAC Plot for level 2 data

Table 14.3 Loadings a_{i1}, a_{i2}, using concordance method for level 2 indicators

	LOV_1	LOV_2
BUF	0.380952	0.571429
IR	0.761905	0.333333
BA	0.428571	0.380952
INV	0.619048	0.476191
SHA	0.571429	0.476191
SS	0.333333	0.428571
FPWL	0.190476	0.476191

$$I_i = a_{i1}{}^*\mathrm{LOV}_1 + a_{i2}{}^*\mathrm{LOV}_2$$

We define

$$g_i = (a_{i1} + a_{i2})/\Sigma(a_{i1} + a_{i2}) \tag{14.5}$$

The quantity g_i is the final weight with which we combine the (normalized) I_i to a data-driven composite indicator (DDI):

$$\mathrm{DDI} = \Sigma g_i{}^* I_i \tag{14.6}$$

Looking at the data-based weights from POSAC (Table 14.4), we see that the data-based index gives approximately the same amount of weight to all the indicators with a little less weight to stream stressor and FPWL stressor. When compared to the investigator-based SWR index, the POSAC-derived weights give less weight to the indicators IR, SHA, and SS than does SWR, since SWR gives all three of these indicators a weight of 0.25.

Table 14.4 Weights for level 2 indicators

Indicator	DDI (POSAC)	SWR
BUF	0.150	0.063
IR	0.154	0.250
BA	0.137	0.063
INV	0.169	0.063
SHA	0.159	0.250
SS	0.126	0.250
FPWL	0.106	0.063

14.8 Summary and Commentary

Considerable effort is expended in the assessment of the quality of watersheds. Three levels of indicators are defined, where level 1 indicators are the cheapest and the level 3 indicators the most expensive.

We performed some partial order analyses to see how the wetlands can be compared under the abstract principle of "environmental health," without applying a composite indicator. In doing this, we study the system of order relations and consequently it is of interest to see how this system of order relations, displayed in Hasse diagrams, changes if we delete indicators from the data matrix. We found which indicators are important and saw that IMP (of the level 1 indicators) is very important, whereas the sensitivity values of the indicators of level 2 are more spread out, albeit INV turned out to be the most important one.

A large part of this chapter is devoted to compare the different systems. As a main result, we see that there is some indication that the level 1 indicators seem to be better proxies for level 3 than level 2 indicators.

As far as the process of constructing composite indicators is considered, the POSAC method may be a good alternative and it may be an issue to discuss the weights of SWR because the order theoretical approach found a different set of weights.

References

Brooks, R.P., McKinney-Easterling, M., Brinson, M., Rheinhardt, R., Havens, K., O'Brien, D., Bishop, J., Rubbo, J., Armstrong, B. and Hite, J. (2007). A Stream–Wetland–Riparian (SWR) index for assessing condition of aquatic ecosystems in small watersheds along the Atlantic Slope of the eastern U.S. (Manuscript).

Patil, G.P. (2001). Penn state environmetrics and ecometrics, cross-disciplinary classroom notes. University Park, PA: Center for statistical Ecology and Environmental Statistics, Penn State University.