

# Increasing the Efficiency of Transboundary Water Management: A Regionalization Approach

Dirk Huchtemann<sup>1</sup>, Manuel Frondel<sup>2</sup>

<sup>1</sup>*Forschungszentrum Juelich GmbH, Juelich, Germany*

<sup>2</sup>*Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Essen, and Ruhr-Universität Bochum, Bochum, Germany*

*E-mail: d.huchtemann@fz-juelich.de, frondel@rwi-essen.de*

*Received December 18, 2009; revised April 23, 2010; accepted May 6, 2010*

## Abstract

In this paper, we develop a methodological approach to increase the efficiency of watershed management approaches that focuses on the reduction of diffuse water pollution. The basic idea is that the efficiency may be improved by identifying and constituting homogenous groups of contiguous administrative units of a watershed, which jointly implement water pollution reduction measures. Homogeneity means similarity of group members with respect to a set of local and environmental characteristics, such as the degree of pollution and abatement costs. We empirically apply our methodology to the watershed of the German river Ems and identify homogenous groups of contiguous administrative units using cluster analysis methods implemented in a Geographical Information System.

**Keywords:** Diffuse Water Pollution, Environmental Federalism, Cluster Analysis, EU Water Framework Directive

## 1. Introduction

Nitrate and phosphorus entry into surface water mainly originates from mineral and organic fertilizers employed in the agricultural sector. This sector is responsible for the vast majority of diffuse—or non-point-source—water pollution in Germany. Within the period 1993 to 1997, the average contribution of the agricultural sector to the overall nitrate and phosphorus entry from non-point sources was as large as 94.1% and 83.7%, respectively [1]. In contrast to pollutants originating from point sources, such as sewage disposal systems and livestock facilities, pollutants from non-point sources, such as agricultural drainage, are difficult to combat.

Diffuse water pollution disperses, thereby affecting less polluted areas and, specifically, potable water resources. Nitrate and phosphorus entries, the two most significant water pollutants, lead to the turbidity and malodor of surface water, as well as its accelerated eutrophication. Given all these undesirable effects, the further reduction of diffuse pollution is a major goal of the EU Water Framework Directive 2000/60/EC (Article 10), although the level of diffuse pollution has been significantly reduced during the past twenty years [2].

Article 3 of this directive demands the coordination of water management within river basin districts, which usually belong to several administrative units such as federal states. As a result, watersheds are affected by frequently distinct and simultaneous actions of several administrations, and possible deficiencies in coordination may create inefficiencies [3].

In this paper, we develop a methodological approach to improve the economic efficiency in the reduction of diffuse water pollution. It rests on the identification and formation of homogenous groups of contiguous administrative units of a watershed. By jointly implementing pollution reduction measures, these homogenous groups are able to diminish negative spillover effects and externalities. Homogeneity refers to characteristics such as willingness-to-pay for pollution reduction, abatement costs, and the degree of pollution, which varies with, e.g., agricultural production.

Our methodology proposed, commonly referred to as a *regionalization approach* [4], is exemplified here by an application to the German side of the watershed of the river Ems. To identify homogenous groups of administrative units within this watershed, we use cluster analysis methods implemented in a Geographical Information System (GIS). As the implementation of joint pollution mitigation

measures is only sensible and manageable in contiguous areas, the spatial relationship among administrative units is an essential variable for this cluster analysis.

In the following section, we explain our regionalization approach. Section 3 provides a concise description of our sample region. Empirical results for our sample region are presented in Section 4. The last section summarizes our findings and provides conclusions.

## 2. Methodological Framework

To avoid negative spill-over effects and inefficiencies due to the divergence between natural watershed areas and districts, the prevailing administration units, we suggest designing new artificial administrative structures consisting of contiguous sets of communities that are – in terms of diffuse water pollution, abatement costs, etc. – more homogeneous than the districts. For this purpose, we must address two methodological issues: 1) the definition of a set of homogeneity criteria and 2) the measurement of homogeneity levels.

In our example of the Ems watershed, the German stretch of which covers 227 communities and 22 districts, the set of homogeneity criteria consists of two variables: *abatement cost* and *discharge rates of nitrate*. Assuming that the livestock density in the agricultural sector is limited to only one animal unit per hectare, the abatement cost data originates from an agro-economic forecast model called RAUMIS [5], which was applied within the interdisciplinary research project REGFLUD<sup>1</sup>. This project is also the source for the data on *discharge rates of nitrate*, which was determined using a GIS-based model developed at Forschungszentrum Jülich.

Given that the implementation of joint water pollution mitigation measures is only sensible and manageable for contiguous administrative units, our clustering approach takes particular account of the spatial relationships of the communities (on cluster analysis and other regionalization approaches see [8] and [9]). Our approach is therefore embedded in a Geographic Information System (GIS). To show the importance of taking spatial relationships into account, **Figure 1** displays the outcome of an artificial clustering example when spatial distances among communities are either neglected or accounted for.

This illustrative example, inspired by [10], comprises 16 communities and two variables: discharge rates of nitrate and abatement costs. If spatial relationships are ignored, the resulting clusters, called classes, are usually not contiguous and exhibit conjoint borderlines only by chance, as indicated by the left-hand side of the lower

part of **Figure 1**. In contrast, taking account of spatial relationships leads to so-called regions, that is, clusters that exclusively consist of contiguous administrative units, as shown on the right-hand side of the lower part of **Figure 1**.

In contrast to e.g. [11], who employ a contiguity matrix to capture spatial relationships, we proceed by determining artificial epicenters for each unit – called centroids – and including their x- and y-coordinates to our cluster analysis. The advantage of our method compared to approaches using a contiguity matrix is the parsimonious number of variables: We only need to add two additional variables, CX and CY, indicating the x- and y-positions of all centroids, rather than a matrix containing the information about spatial relationships among all administrative units, which, in practice, requires hundreds of variables [11].

In short, we are left with merely four variables – abatement costs, discharge rates of nitrate, CX, and CY – for our cluster analysis, rather than several hundred variables. For operational convenience, we simply assign equal weights to these four variables. Any unequal weighting scheme – for instance, by assigning lower weights to the geographical variables CX, and CY because the other variables are regarded as more important – would be arbitrary.

Additionally, we need a measure that helps us to decide whether two communities, say,  $p$  and  $q$ , are similar to each other regarding the set of selected variables. Among a variety of such measures (see [12]), we choose the *Squared Euclidian metric*:

$$d_{pq} = \sum_{j=1}^N (x_j^p - x_j^q)^2, \quad (1)$$

with  $x_j^p$  and  $x_j^q$  denoting the concrete realizations of variable  $X_j$  for community  $p$  and  $q$ , respectively. Before calculating the Squared Euclidian distances  $d_{pq}$ , one has to standardize the variables. Upon standardization, each of the transformed variables displays zero mean and a uniform standard deviation of 1.

The more similar two communities are the closer to zero is the Squared Euclidian metric. Eventually, we come up with a symmetric  $(227 \times 227)$  matrix  $(d_{pq})$ , representing the basis for the clustering procedure. We then need to choose among several clustering algorithms that help us to identify the ultimate number of clusters such that the heterogeneity across, and the homogeneity within, all clusters is maximized. For this purpose, we have chosen the centroid method<sup>2</sup> (see [13] for other methods).

The second methodological issue refers to the measurement of the homogeneity of our cluster analysis out

<sup>1</sup>The overall objective of REGFLUD was the development and application of multi-criteria scientific methods to set up a Decision Support System aimed at reducing nonpoint source pollution in river catchments subject to economic feasibility and social acceptability. The project was funded by the Federal Ministry of Education and Research (BMBF) [6,7].

<sup>2</sup>The centroids used in this clustering algorithm must not be confused with geographic centroids: These centroid vectors include the coordinates of geographic centroids, but also the respective values of the variables *abatement cost* and *discharge rates of nitrate*.

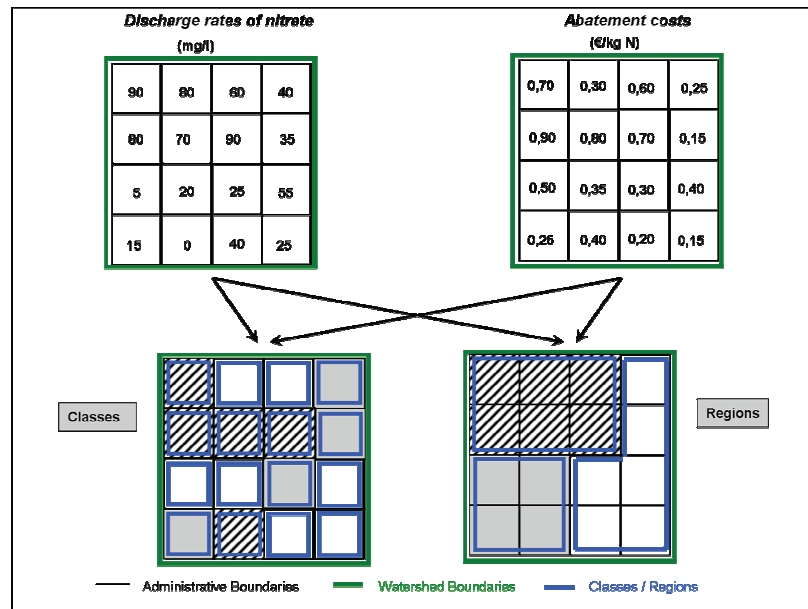


Figure 1. Clustering with and without incorporating spatial distances of communities.

comes — the set of output clusters. Commonly, homogeneity is examined for each criteria variable  $X_j$  individually. Homogeneity tests are based on a statistical measure defined as the ratio of the within-cluster variance  $Var^G(X_j)$  of a given cluster  $G$  and the between-cluster variance  $Var(X_j)$ :

$$F = \frac{Var^G(X_j)}{Var(X_j)}. \quad (2)$$

The letter  $F$  indicates that the probability distribution of this test statistic is the  $F$ -distribution. While the  $F$ -value of a cluster that is perfectly homogeneous with respect to criteria variable  $X_j$  is zero, such desirable values do not occur in practice. In contrast, an  $F$  value equaling, or even exceeding, unity is not desirable. In this case, the within-cluster variance equals, or is even larger than, the respective between-cluster variance, which contradicts the inherent idea of clustering. Of course, the preferred result of any clustering approach is that all the clusters'  $F$ -values are substantially smaller than unity for *all* criteria variables. In this case, the clusters are called *totally homogeneous*.

For interpretational purposes, another test statistic is frequently employed. It is based on the relative difference of the within-cluster means  $\bar{x}_j^G$  for a given cluster  $G$  and  $\bar{x}_j$ , which is the sample mean of criteria variable  $X_j$ :

$$t = \frac{\bar{x}_j^G - \bar{x}_j}{std(X_j)}, \quad (3)$$

with  $std(X_j)$  denoting the sample standard deviation of variable  $X_j$ . This test statistic, which follows Student's  $t$  distribution, reflects the relative deviation of a cluster  $G$

from the respective sample means. A large  $t$ -value indicates that a cluster  $G$  differs strongly from the sample average of variable  $X_j$ .

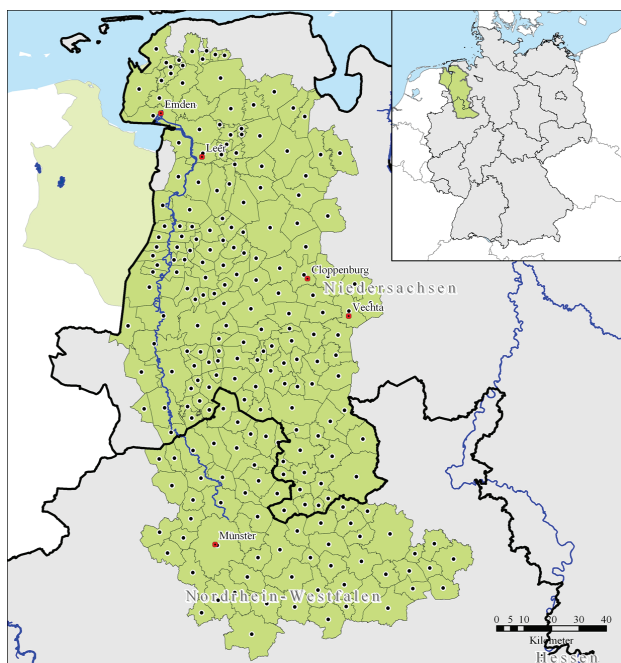
### 3. Sample Region

Our approach is applied to the German side of the watershed of the river Ems, the region displayed in green in **Figure 2**. The major part of this watershed is located in the North-West of Germany and stretches across the boundaries of two federal states, North-Rhine Westphalia and Lower Saxony. 227 German communities are either completely located in this watershed or share a significant part of it. The Dutch side of the watershed - displayed in light green in **Figure 2** - must be ignored in our application due to the lack of data.

The Ems watershed covers a wide range of different landscapes, exhibiting different hydrological, hydrogeological, and socio-economic characteristics. Agricultural production, particularly livestock farming, dominates the north-eastern part of this region, which is characterized by meagre, sandy soils. The dominance of livestock farming causes an abundance of organic fertilizers in the form of liquid manure. Generally, this liquid manure, commonly distributed to agricultural acreage, leads to nitrate entry into surface water and, hence, diffuse water pollution. Its degree depends upon the type of soil and the varying nitrate retention and degradation ability.

### 4. Regionalization of the Ems Watershed

In this section, we apply our regionalization approach to



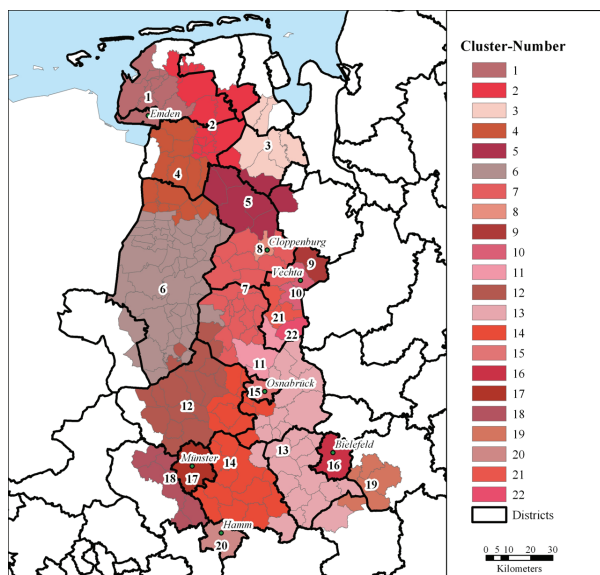
**Figure 2.** The Ems watershed and its communities in Germany.

the communities of the German side of the watershed of the river Ems. Our aim is to identify contiguous clusters among the 227 communities that are more homogeneous with respect to *abatement costs* and *discharge rates of nitrate* than the prevailing administrative units (districts). Using the method described in Section 2, we end up with 22 clusters, which is only by chance equal to the number of districts. All outcome clusters displayed in **Figure 3** are indeed contiguous groups of communities and most of these clusters are substantially different from the prevailing districts, although there are some similarities.

As a major result, the incongruity of district and cluster borderlines shown in **Figure 3** indicates that the homogeneity among communities with respect to abatement cost and discharge rates of nitrate might be improved by re-grouping these communities along the lines of the outcome of our cluster analysis. F-values reported in **Table 1** are well below the critical value of 1 for all clusters, indicating that the clustering outcome is totally homogenous with respect to the variables *discharge rates of nitrate* and *abatement costs*.

It is somewhat surprising that, except for three heterogeneous groups of communities, the prevailing administrative structure given by the districts of the Ems watershed appears to be quite homogeneous.

The *t*-values reported in **Table 1** indicate that the abatement costs of clusters 8, 9, 10, 15, 16, 17, 21, and 22 are significantly larger than on average. With the exception of clusters 8 and 16, these clusters also display larger nitrate discharge rates than the sample average. Yet, there is no perfect positive correlation between nitrate



**Figure 3.** The outcome of a regionalization approach: Clusters of the Ems watershed.

entries and abatement cost, since the clusters 1, 4, 12, 13, 14, and 18 exhibit discharge rates of nitrate that are above the sample average, whereas the abatement costs of these clusters are below the average.

## 5. Summary and Conclusions

Current watershed management practices in Germany, as well as in other European countries, usually do not take particular account of the hydrological boundaries of watersheds. Yet, the fact that international and interregional watersheds are affected by the frequently distinct but simultaneous actions of several administrative units, such as districts, typically creates inefficiencies. Identifying and clustering adjacent communities of watersheds that are more homogeneous groups than the prevailing districts and that will then jointly implement pollution reduction measures has been suggested in this paper as a way to potentially increase the efficiency of watershed management.

Using cluster analysis methods, our regionalization approach is perfectly in line with Article 3 of the EU Water Framework Directive 2000/60/EC, which demands the coordination of the water management within river basin districts. Our approach is based on the theory of environmental federalism that deals with the efficient assignment of federal competencies – see, for example, [14]. Constitutional elements of this theory are the *principles of subsidiarity* and *fiscal equivalence* [15,16]. The principle of subsidiarity implies that competencies ought to be assigned to the lowest federal level that suffices to meet a certain task. The efficient provision of, specifically, the public good “clean water” requires the matching of the natural areas of watersheds with administra-

**Table 1. Regionalization outcome.**

Cluster	Number of Communities	F-values abatement cost	t-values abatement cost	F-values nitrate	t-values nitrate
1	12	0.009	-0.491	0.119	0.415
2	19	0.006	-0.426	0.285	-1.138
3	6	0.004	-0.245	0.239	-0.538
4	18	0.021	-0.297	0.422	1.235
5	7	0.832	0.808	0.733	1.066
6	51	0.036	-0.312	0.295	-0.897
7	22	0.298	0.078	0.172	-0.324
8	1	-	3.510	-	-0.495
9	2	0.001	2.214	0.005	1.139
10	2	0.440	7.285	0.032	0.699
11	3	0.795	0.886	0.222	0.987
12	18	0.062	-0.148	0.102	0.107
13	27	0.064	-0.149	0.179	0.231
14	21	0.040	-0.199	0.366	0.999
15	1	-	4.378	-	0.791
16	1	-	1.587	-	-0.764
17	1	-	1.341	-	0.222
18	6	0.003	-0.223	0.128	1.926
19	6	0.024	-0.245	0.662	-1.088
20	1	-	-0.413	-	-0.846
21	3	0.716	2.279	0.049	0.492
22	1	-	4.117	-	1.896

Note that for degenerated “clusters” including only one community *F*-values are undefined.

tive territories, and is in accordance with the principle of fiscal equivalence, which is also called perfect mapping.

According to [17], there are three criteria characterizing an ideal administrative unit: 1) Preferences for a public good, such as clean water, need to be homogenous among inhabitants of such a unit; 2) costs for the provision of this public good must be similar within the administrative unit, and 3) specific policy measures aimed at improving the provision of the public good must *only* affect this unit’s area (*ideal geographical scope*).

Along the lines of these ideas, we have developed a regionalization approach that has been empirically applied to the communities belonging to the German side of the watershed of the river Ems. Since the implementation of joint pollution mitigation measures is only sensible and manageable in adjacent areas, it is critical to cluster this watershed’s communities in such a way that the resulting clusters build contiguous areas. Hence, the spatial relationship among administrative units is an essential variable for any cluster analysis employed to identify more homogeneous clusters of communities than the prevailing districts within this watershed. We therefore use cluster analysis methods implemented in a Geographical Information System (GIS). Homogeneity means that communities are similar regarding a defined set of variables, such as the degree of pollution and

abatement costs for water pollution reduction.

The results based on empirical data on abatement costs and average nitrate discharge rates of our sample communities demonstrate that the suggested methodological approach allows for creating regions that are more homogeneous than the prevailing districts, *i.e.*, the current administrative units. The resulting regionalization, however, represents only one possible spatial structure. Further research has to concentrate on determining and comparing the magnitude of the efficiency gains from different solutions.

Moreover, such reorganization might be problematic, because it might interfere with the current administrative structure and therefore may cause substantial legal problems that, ultimately, may hamper the desired watershed cooperation [18]. Nevertheless, the results of our regionalization approach may be useful when attempting to implement the EU Water Framework Directive.

## 6. Acknowledgments

We are grateful for invaluable comments and suggestions by Christoph M. Schmidt and Colin Vance. We also gratefully acknowledge funding by the Federal Ministry of Education and Research (BMBF) within the framework of the REGFLUD project.

## 7. References

- [1] FME, "Water Resources Management: Emissions into Surface Waters and the Sea," Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, Berlin, 2001, pp. 49-51.
- [2] FME, *ibid.*, p. 58.
- [3] T. Moss and H. J. Kujath, "Chancen und Grenzen des Institutionenwandels," *Nachhaltige Entwicklung in der Wasserwirtschaft: Konzepte, Planung und Entscheidungsfindung*, Schriftenreihe Wasserforschung, B. Weigert and C. Steinberg, Eds., Berlin, Vol. 7, No. 1, 2001, pp. 305-324.
- [4] R. J. Johnston, D. Gregory and D. M. Smith, Eds., "The Dictionary of Human Geography," 3rd Edition, Oxford, 1997, pp. 71-87.
- [5] H. Gömann, P. Kreins, R. Kunkel and F. Wendland, "Koppelung agrarökonomischer und Hydrologischer Modelle," STE-Reprint 29/2003, Jülich, 2003.
- [6] P. Kreins, H. Becker, D. Hecht, D. Huchtemann and F. Wendland, "Management regionaler Flusseinzugsgebiete in Deutschland (REGFLUG) – Rahmenbedingungen und Politikoptionen bei diffusen Nährstoffeinträgen (n und P) der Landwirtschaft in den Rhein und die Ems," *Interdisziplinäre Methoden des Flussgebietmanagements*, J. Möltgen and D. Petry, Eds., IfGI Prints 21, University of Münster, Münster, 2004.
- [7] <http://www.faa-bonn.de/reghome.htm>
- [8] B. S. Everitt, S. Landau and M. Leese, "Cluster Analysis," 4th Edition, Oxford University Press, Oxford, 2001.
- [9] J. C. Duque, R. Ramos and J. Surinach, "Supervised Regionalization Methods: A Survey," *International Regional Science Review*, Vol. 30, No. 3, 2007, pp. 195-220.
- [10] G. Urfei, "Agrarumweltpolitik nach den Prinzipien der Ökonomischen Theorie des Föderalismus – Ein Regionalisierungsansatz zur Territorialen Abgrenzung Effizienter Politikaktionsräume," Dunker und Humblot, Berlin, 1999.
- [11] D. Martin, "Developing the Automated Zoning Procedures to Reconcile Incompatible Zoning Systems," *Proceedings of the 6<sup>th</sup> International Conference on Geo-Computation 2001*, University of Queensland, Brisbane, 2001.
- [12] B. S. Everitt, S. Landau and M. Leese, *ibid.*, pp. 35-45.
- [13] B. S. Everitt, S. Landau and M. Leese, *ibid.*, pp. 55-89.
- [14] W. E. Oates, "A Reconsideration of Environmental Federalism," *Recent Advances in Environmental Economics*, J. A. List and A. J. DeZeeuw, Eds., Edward Elgar, Cheltenham, 2002.
- [15] A. Breton, "A Theory of Government Grants," *Canadian Journal of Economics and Political Science*, Vol. 31, No. 2, 1965, pp. 175-187.
- [16] J. Buchanan, "An Economic Theory of the Clubs," *Economica*, Vol. 34, No. 3, 1965, pp. 1-14.
- [17] M. Olson, "The Principle of 'Fiscal Equivalence': The Division of Responsibilities among Different Levels of Government," *American Economic Review*, Vol. 59, No. 2, 1969, pp. 479-487.
- [18] T. Moss, "Raumwissenschaftliche Perspektiverweiterung zur Umsetzung der EU-Wasserrahmenrichtlinie, Das Flussgebiet als Handlungsraum," *Institutionenwandel durch die EU-Wasserrahmenrichtlinie aus raumwissenschaftlichen Perspektiven*, *Stadt und Regionalwissenschaften*, Bd. 3, Lit Verlag, Münster, 2003.