A Novel EOR Screening Approach based on Bayesian Clustering and Principal Component Analysis

Martina Siena[1], Alberto Guadagnini[1,2], Ernesto Della Rossa[3], Andrea Lamberti[3], Franco Masserano[3], Marco Rotondi[3]

[1]Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Piazza L. Da Vinci 32, 20133 Milano, Italy

[2]Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, 85721, USA

[3]eni S.p.A. -Via Emilia 1, 20097 San Donato Milanese (Milano) Italy

Abstract

We present and test a new screening methodology to discriminate amongst alternative and competing Enhanced Oil Recovery (EOR) techniques to be considered for a given reservoir. Our work is motivated by the observation that, even if a considerable variety of EOR techniques have been successfully applied to extend oilfield production and lifetime, an EOR project requires extensive laboratory and pilot tests prior to field-wide implementation and preliminary assessment of EOR potential in a reservoir is critical in the decision-making process. Since similar EOR techniques may be successful in fields sharing some global features, as basic discrimination criteria we consider fluid (density and viscosity) and reservoir formation (porosity, permeability, depth and temperature) properties. Our approach is observation-driven and grounded on an exhaustive data-base which we compile upon considering worldwide EOR field experiences. A preliminary reduction of the dimensionality of the parameter space over which EOR projects are classified is accomplished through Principal Component Analysis (PCA). A screening of target analogs is then obtained by classification of documented EOR projects through a Bayesian clustering algorithm. Considering the cluster which comprises the EOR field under evaluation, an inter-cluster refinement is then accomplished by ordering cluster components on the basis of a weighted Euclidean distance from the target field in the (multidimensional) parameter space. Distinctive features of our methodology are that (a) all screening analyses are performed on the database projected onto the space of principal components, and (b) the fraction of variance associated with each principal component is taken as weight of the Euclidean distance we determine. As a test bed, we apply our approach on three fields operated by eni. These include light, medium and heavy-oil reservoirs, where Gas, Chemical and Thermal EOR projects have been respectively proposed. Our results are (a) conducive to the compilation

of a broad and extensively usable data-base of EOR settings and (b) consistent with the field observations related to the three tested and already planned/implemented EOR methodologies, thus demonstrating the effectiveness of our approach.

1. Introduction

Enhanced Oil Recovery (EOR) techniques that have been successfully applied to date can be grouped into three main categories: (i) thermal methods, (ii) gas and water-alternating-gas (WAG) injection, and (iii) chemical injection. Regardless of the technique, the route leading from design to full field-scale implementation of an EOR project is always complex and fraught with challenges.

Careful and detailed preliminary studies must be performed to reduce uncertainty and minimize the risk of failure. These analyses comprise laboratory tests, and progress through reservoir characterization and simulation, design and implementation of pilot tests, to the final design and implementation of the full field project. Moreover all the above mentioned phases involve investments that can be risky if not properly supported by a preliminary cost efficient screening phase. Hence, a key element in the decision-making process is, first and foremost, the assessment of the EOR potential for a target reservoir. This is the critical goal accomplished by the practice of EOR screening, which is meant to provide a first metric to be employed for risk reduction with modest capital investment. The main problem addressed by screening is then the assessment of the most suitable EOR technique for a target reservoir.

All existing screening methods typically require: (i) a knowledge of target field characteristics, and behavior, (ii) a detailed knowledge of previous EOR experiences, and (iii) a thorough understanding of the recovery mechanisms upon which EOR techniques are based. Mechanisms driving oil displacement are a major distinctive feature amongst EOR techniques. The effectiveness of a given mechanism is strictly related to the properties of the reservoir and the oil. These properties are therefore critical elements for the assessment of the potential of different and sometimes competing EOR techniques. A total of six parameters are typically considered in the literature as representative of a given setting, i.e., reservoir porosity, permeability, depth and temperature and oil density and viscosity (Alvarado et al., 2002, Babushkina et al., 2013, Kamari et al., 2010). Adoption of these parameters is based on correlation studies (e.g., Alvarado et al., 2002) aimed at reducing redundancy of information as well as on practical considerations, which are related to the relative ease with which these quantities can be evaluated in the field.

There are two basic categories of screening approaches: conventional and advanced methods. Conventional methods (Taber et al., 1997a,b, Al Adasani and Bai, 2011) tend to regard a target EOR technique as potentially effective on a given field if the field oil and reservoir parameters are comprised within prescribed ranges, estimated on the basis of expert opinion and information on previous projects. As an example, Fig. 1 illustrates the conventional screening criteria presented by Taber et al. (1997a, b). Regions enclosed within the ellipses constitute qualitative representations of the conditions of applicability of the main EOR techniques considering only reservoir depth and oil viscosity. The plot suggests that deep (i.e., 4000 - 18000 ft) light-oil bearing (viscosities in the range of 0.01 to 10 cp) reservoirs may be exploited by means of gas injection EOR. Otherwise, shallow (i.e., 100 - 4000 ft) heavy-oil bearing (viscosities in the range of 10 to 100000 cp) reservoirs may be developed by thermal projects. Instead, chemical EOR projects are generally applied under intermediate conditions which are typical of reservoir depths in the range of 1000-9000 ft and oil viscosities in the range of 1 to 100 cp. As highlighted by Alvarado and Manrique (2010), the notable limitation of conventional methods is that they only provide "go/no go" criteria, without any additional detailed information about EOR strategies developed in similar reservoirs.

Advanced methods usually aim at identifying successful EOR applications on reservoirs that share some analogies (in terms of reservoir and oil properties) with the target field under scrutiny. Recalling that diverse EOR techniques are based upon different displacement mechanisms and that the effectiveness of each mechanism depends on the physical properties of the field, the underlying concept is that similar EOR techniques should be successfully applied to fields sharing similar features. Hence, advanced EOR screening methods address the way analogy can be quantified and employed as an EOR screening criterion.

Advanced methods tackle the problem of EOR screening in the framework of multi-dimensional statistical analysis. These approaches rely on several data mining and artificial intelligence strategies, including neural networks (Surguchev and Li, 2000; Kamari et al., 2014), expert systems (Gharbi, 2000, Abbas and Song, 2011), fuzzy inference (Anikin, 2014), or Bayesian Networks (Moreno et al., 2014; Zerafat et al., 2011).

The advanced method presented in details by Alvarado and Manrique (2010) and previously documented by Alvarado et al. (2002), Manrique et al. (2003), and Manrique et al. (2009), is basically a representation in the form of two-dimensional expert maps of a database of successful EOR projects, indexed by the set of the six parameters listed above. These maps are obtained through a multidimensional data projection and allow qualitative identification reservoir clusters of EOR projects by visual inspection. Each identified cluster is then regarded as representing a pseudo-typology of reservoir and all fields belonging to the same cluster are considered to be similar. The screening strategy developed by Trujillo et al. (2010) combines a conventional approach, based on the property ranges suggested by Taber et al. (1997a,b) with an advanced approach, in which analogy/similarity is defined by ranking fields included in the available database according to a similarity score. The latter is computed as the average of the differences (rescaled to the unit support [0,1]) between fluid/reservoir properties of the database and target fields. In the method illustrated by Moreno et al. (2014) the search of analogs does not rely solely on average properties of oil and reservoir, but also on a more detailed characterization of the target field heterogeneity, using available three-dimensional numerical model. With this approach, it is also possible to provide predictions of oil displacement and production associated with a given EOR technique by simulating the selected development scenario.

Babushkina et al. (2013) define and investigate analogy by applying a k-means clustering algorithm on the six-dimensional space of oil and reservoir properties. The EOR potential of a target field is estimated by interpolation of the recovery factors associated with the (eventually different) EOR techniques of projects belonging to the same cluster.

In this work, we introduce a new advanced screening methodology which relies on the combination of two modern approaches for analogy assessment. These are respectively based on (a) the grouping of fields into clusters through Bayesian hierarchical clustering, and (b) the computation of suitable Euclidean distances between fields belonging to the same cluster. These approaches are highly compatible and conducive to an effective screening. The selected Bayesian hierarchical clustering algorithm enables one to deconstruct the available database of observed field projects and form groups of analogous fields according to a rigorous probabilistic criterion. This approach recognizes that all elements assigned to a given group after cluster decomposition have the same degree of analogy. As such, it has the benefit of circumventing the ambiguities which are typical of distance-based methods and related to the definition of the largest distance beyond which analogy should not be considered meaningful. Refinement of the identification of analogs is then accomplished by assigning a rigorous and consistent measure of similarity to each cluster element and ranking each of these elements/fields according to this metric.

A principal component analysis (PCA) is performed on the elements of the available field case database to reduce redundancy of information associated with cross-correlation between the selected properties.

Additional elements of novelty of our screening approach stem directly from the application of PCA, which serves not only for visualization purpose and dimensionality reduction (as, e.g., in Babushkina et al., 2013). All screening analyses in our methodology are performed on the database projected onto the space of principal components. PCA also plays a key role in the choice of the inter-cluster distance metric, as the fraction of variance associated with each principal component is used as a weight to the Euclidean distance between cluster elements.

The work is structured as follows. Section 2 describes the reference database we consider and the methodological workflow, focusing on the operations performed on the database at each step. Section 3 details the blind testing of our screening procedure on three reservoirs operated by eni where the successful use of given EOR techniques is documented. Conclusion and remarks are then summarized in Section 4.

2. Materials and Methods

Figure 2 provides a graphical depiction of the workflow associated with our screening procedure. We structure our approach onto three main stages, respectively aimed at (i) building a database of EOR projects implemented in the field; (ii) identifying within the database analogs to the target field where EOR is envisioned to be implemented; and (iii) analyzing the results in terms of relative benefits of EOR techniques applicable to the target field. The first stage relies on an extensive review of successful pilot/full-field applications of EOR projects. We collect extensive information mainly from SPE journal publications and conference proceedings, Oil and Gas Journal biennial surveys (Koottungal, 2008, 2010, 2012, 2014) and annual reports published and made available by various oil companies. This first stage enabled us to compile a rich database comprising of 250 worldwide cases, where applications of diverse EOR techniques, ranging from thermal, chemical and gas/WAG injection strategies have been documented. Thermal approaches considered in the database include in-situ-combustion, hot-water and steam-injection projects. Chemical EOR projects are either based on polymer, surfactant or Alkali-Surfactant-Polymer (ASP) flooding. Gas/WAG injection methods differ in terms of (i) the type of injected gas (Hydrocarbons, Carbon Dioxide or Nitrogen); (ii) miscibility/immiscibility of the injected gas with reservoir fluid. Figure 3 displays the relative occurrence of EOR techniques within the database. The proportions of the diverse EOR techniques associated with our collection of data are consistent with those associated with data sets analyzed in previous works (see, e.g., Fig. 2 of Al Adasani and Bai, 2011; or Fig. 1 of Moreno et al., 2014). Figure 4 depicts the location on a world map of all cases included in our database. This plot illustrates the density of EOR techniques per geographic area and can serve as a preliminary qualitative screening tool. Each EOR project in the database is indexed by the set of six field parameters listed in Section 1, i.e., average reservoir rock porosity, ϕ; permeability, k; depth, D; temperature, T; and oil viscosity, μ, and density, API. The set of parameters associated with each target reservoir is embedded into the database. Ranges of variation of all parameters are listed in Table 1.

2.1. Data Pre-processing.

Our data pre-processing relies on a sequence of preliminary transformations that are applied to the database to assist the screening procedure. These transformation are robust standardization (RS) and principal component analysis (PCA), which are applied in sequence to the data, as detailed in the following.

2.1.1. Robust Standardization (RS).

The six selected properties are characterized by remarkably different units and ranges of variation. As permeability and viscosity may vary over several orders of magnitudes, we consider a logarithmic transformation for our analysis. Data are then subject to standardization. In its common form, this transformation centers and rescales data respectively by their sample mean and standard deviation. To prevent the effect of outliers, robust standardization is here preferred (Daszykowski et al., 2007). In this context, data are centered by their median and rescaled by their median absolute deviation (MAD), i.e., robust-standardized value ξ_{RS} of a given variable ξ in our database is

$$\xi_{RS} = \frac{\xi \text{ - median}(\xi)}{\text{MAD}(\xi)}; \qquad \text{MAD}(\xi) = 1.4826 \cdot \text{median}\left(\left|\xi - \text{median}(\xi)\right|\right) \tag{1}$$

2.1.2. Principal Component Analysis (PCA).

A substantial simplification of the screening problem can be obtained by noting that some of the selected variables are cross-correlated. This is clearly visible in some of the diagrams appearing in the lower part of the matrix plot of Fig. 5, which depicts scatterplots between pairs of RS-variables in our database. The redundancy of information associated with cross-correlation can be effectively removed by applying a PCA to our database. Essentially, PCA allows mapping the system from the space of RS-variables to a new space of uncorrelated variables (i.e., the principal components, PCs). Each of the PCs is a linear combination of the original properties and carries information about a fraction of the multivariate overall variability of the system. The effect of PCA in removing cross-correlation is highlighted in the above-diagonal diagrams depicted in Fig. 5, representing scatterplots between pairs of PCs. As a complement to these results, Fig. 6 depicts vectors indicating the contribution of each of the RS-variables to the first two PCs. Figure 7 illustrates the fraction of the overall system variance that is explained by each PC. These results clearly show that the information content embedded in the first three PCs explains 95% of the overall system variability. Hence, we choose to perform all screening operations described in the following in the three-dimensional space identified by these PCs, as a tradeoff between completeness of information and level of complexity associated with the interpretation of results.

2.2. Bayesian Hierarchical Clustering Algorithm.

We start our analogy assessment by employing a clustering algorithm. Clustering is commonly applied in classification problems because it allows organizing elements into disjoint groups according to prescribed rules that depend on the algorithm of choice. Here, we select a Bayesian Hierarchical Clustering (BHC) algorithm (Heller and Ghahramani, 2005). As other hierarchical methods, this approach is bottom-up agglomerative. Starting from a setting where each element forms a cluster, the algorithm iteratively merges the two most similar clusters. As opposed to traditional clustering methods based on distance-based criteria, the notion of similarity is inferred in BHC in a probabilistic context, by considering the Bayesian probability that the two clusters should be merged into a unique cluster. This probability is evaluated considering the basic assumptions behind BHC, according to which (a) the whole database is sampled from a mixture associated with a given distribution model, and (b) elements of the same cluster are sampled from the same component contributing to the mixture, each of these components being characterized by a given set of parameters. Hence, at each step the algorithm merges the two clusters whose elements have the highest probability of being sampled by the same set of distribution parameters. We refer to Heller and Ghahramani (2005) and Sirinukunwattana et al. (2013) for the detailed formulation on the computation of these probabilities. As shown by these authors, the Bayesian approach enables to overcome key limitations which are typically associated with distance-based clustering algorithms (including, e.g., the k-means approach adopted by Babushkina et al., 2013), which are conditioned by the choice of the distance metrics and require to define a priori the final number of clusters. BHC methods allow inferring naturally the final number of cluster, because merging of clusters takes place only if the associated Bayesian probability is larger than 0.5. The merging process stops when all remaining cluster pairs are associated with a probability of being merged which is below this threshold.

2.3. Weighted Euclidean Distance.

In this step of the screening procedure we focus on those clusters containing one or multiple target fields where we wish to provide EOR screening. The inter-cluster distance between group members of the available database and targets is used as a similarity measure. As mentioned in Section 2.1, the distance is evaluated in the space of the first three principal components, $\{\mathrm{PC}\ 1, \mathrm{PC}\ 2, \mathrm{PC}\ 3\}$, through the weighted Euclidean metric

$$\mathrm{dist}\,(\mathbf{x},\mathbf{y}) = \sqrt{\sum_{i=1}^{3} w_i (x_i - y_i)^2} \tag{2}$$

where $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$ are the position vectors of two given points in the space {PC 1, PC 2, PC 3} and w_i are metric weights. The latter have been set as the fraction of total variance associated with each PC

$$w_i = \frac{\mathrm{Var}\,(\mathrm{PC}\,i)}{\sum_{i=1}^{6} \mathrm{Var}\,(\mathrm{PC}\,i)} \tag{3}$$

$\mathrm{Var}\,(\mathrm{PC}\,i)$ being the variance explained by the i-th principal component.

Babushkina et al. (2013) adopted a confidence index to rank EOR methods linked to elements included in the cluster. This index is based on a measure of closeness of the target to each sub-group formed by cluster elements sharing the same EOR technique. Our ranking strategy is different because our metric is based on a quantitative and intrinsic criterion employed to define the distance weights (3) that are typically considered as uniform (e.g., Babushkina et al., 2013) or are determined according to experience. The method of Babushkina et al. (2013) attributes a confidence index to a given EOR technique and not to the specific applications. As a further remark, we recall that a critical element linked to the identification of analogs is the possibility to obtain information about technical details of the associated EOR experiences documented in literature. This element is fully addressed by our method, which provides quantitative measure of the affinity of the target field to each single project.

2.4. Impact of Principal Component Analysis on Bayesian Clustering

We noted that PCA has the advantage of providing quantitative criteria to reduce the dimensionality of the system. This, in turn, leads to a clear graphical representation of the results which can be readily interpreted and used in applications. Fig. 8 depicts a comparison of the results of BHC decomposition as applied to the system of original parameters after robust standardization (Fig. 8a) against those obtained by performing the analysis in the PCs coordinates system (Fig. 8b). Clusters distribution in Fig. 8a is depicted in the phase space of RS-log permeability versus RS-depth. The choice of this two-dimensional space allows for a straightforward comparison with the results obtained with the PCs, because RS-log permeability and RS-depth are the variables with the closest alignment with the directions of the first two Principal Components (see Fig. 6). A close inspection of Fig. 8 suggests that the distinction between diverse clusters is enhanced in the two-dimensional space of our PCs.

As noted above, each PC in our system bears a diverse relative importance. Therefore, it can be argued that performing analog ranking upon relying on a metric defined in the PCs space is more general and would lead to final results which are more rigorous, from a theoretical standpoint. While we adopt this metric for the evaluation of distances, as noted in Sec. 2.3, we do not employ it at the stage of cluster decomposition. Fig. 9a depicts the results of the BHC algorithm applied to the system in the space of Weighted-PCs (WPCs). Joint analysis of Figs. 8b and 9 enables us to note that considering the effects of the weights on the PCs leads to a marked reduction of the ranges of variation of the variables describing our system. This in turn allows for a straightforward identification of a dominant cluster, on the basis of Bayesian criteria. The results depicted in Fig. 9b illustrate the relative fraction of EOR technique applied within this identified cluster and show that the composition of this group is quite heterogeneous. These results suggest that the PC-weights metric would decrease the effectiveness of cluster decomposition as a first screening tool.

3. Application to test cases

3.1. Description of test cases.

We validated our screening method by applying it to three target fields at different stages of deployment, where the effectiveness of specific EOR processes has been already investigated. The complete set of oil and reservoir properties associated with each target is listed in Table 2. The reliability of our methodology is assessed through a blind test comparing the results of our screening procedure against the actual EOR experiences documented or planned for these fields.

Target 1 is a light oil bearing sandstone reservoir in North Africa. Hydrocarbons are located in low net-to-gross ratio fluvial sand units. Production in the field started from day zero with EOR techniques based on crestal miscible hydrocarbon gas injection. The field is also developed by way of peripheral water injection. Following a preliminary study phase, a water injector well was converted to gas injector in 2013 to increase the recovery factor, thus implementing a WAG pilot in the field.

Target 2, is a heavy oil bearing high-permeability shallow sandstone reservoir in South America. Oil viscosity at reservoir condition is approximately 2500 cp. The reservoir is currently produced by natural depletion. Here, the industrial objective is to apply steam injection in the future to increase the recovery factor, which is otherwise very low.

Target 3 is a medium oil viscosity reservoir in North Africa. The reservoir formation is a mixed sandstone/carbonate sedimentary unit, deposited in a marginal to shallow marine environment. Only the sandstone facies in the unit are oil bearing. Oil viscosity at reservoir condition is 23 cp. The very low bubble-point pressure ensures the virtual absence of free gas in the reservoir during production. The initial phase of primary depletion was followed by the implementation of water injection. Following a detailed study and design phase, a polymer injection pilot has been implemented in 2014.

3.2. Screening results.

The application of the Bayesian Hierarchical clustering algorithm described in Section 2.2 to the collection of database elements and targets leads to the identification of 8 clusters. Figure 10 depicts the clusters and targets arrangement in the phase space of the first two principal components, PC 1 and PC 2. As illustrated in the figure, Target 2 belongs to Cluster 1, whereas Targets 1 and 3 belong to Cluster 2. The analyses detailed in the following are focused on these two groups. Figure 11 depicts the results of the screening in terms of percentage of EOR techniques applied within Cluster 1 (Fig. 11a) and Cluster 2 (Fig. 11b). Figure 12 displays distance-based ranking of cluster elements and associated EOR techniques for Targets 1, 2 and 3 (Fig. 12a, 12b and 12c respectively).

3.2.1. Cluster 1 (Target 2).

Figure 11a evidences that the vast majority of the elements in this cluster is associated with thermal methods, in particular with steam injection (83% of the cases) and, to a smaller extent, with hot water injection and combustion (about 5%). Distance-based rankings illustrated in Fig. 12a unambiguously show that the analogs which are closest to Target 2 correspond to steam injection projects, the latter being the technique proposed for this target, as described in Section 3.1. The first two nearest neighbors are both heavy-oil reservoirs, i.e., (a) the Main Tia Juana field, in Venezuela, and (b) the Huanxiling field, in China. Thermal methods played a key role in the production history at (a) since the mid-60s, where 11 Cyclic-Steam-Stimulation (CSS) projects have been implemented from 1964 to 1968, providing a production enhancement of about 5 MSTB over a cumulative production of 39 MSTB (Vega Riveros and Barrios, 2011). A massive implementation of CSS projects throughout the whole Tia Juana field enabled production rate to attain a peak of 160 KSTB/d, followed by a decline to 44 KSTB/d in 1986. The overall recovery factor is estimated to be close to 60%. An up-to-date thermal EOR method named Toe-to-Heel air injection (THAI) has been recently implemented and has been estimated to increase the recovery factor up to 80% (Uribe et al., 2010). Production in the Huanxiling field, in the Liaohe area, started in the 70s and thermal recovery via CSS initiated in 1984. A steam-flooding test started in 1998 and allowed to reach a recovery factor of 30% of the OOIP after 9 years of continuous injection. New steam-flooding projects have been initiated in 2008 (Yuqiu and Yali, 2009).

3.2.2. Cluster 2 (Targets 1 and 3).

Figure 11b depicts the relative fraction of data associated with diverse EOR techniques associated with elements of Cluster 2 and clearly shows that there is a marked dominance of gas/WAG injection EOR applications for the elements of this cluster. About 50% of cases is associated with (continuous or WAG) miscible HC injection projects and approximately 20% of cases with miscible CO2 injection projects. Chemical EOR projects constitute the majority of the remaining fraction, namely polymer (15%) and surfactant (3%) injection.

Target 1.

Figure 12b shows that the cluster elements which are nearest to the target are characterized by the leading EOR technique within the cluster, i.e., miscible HC injection, the only exception being a single case where thermal EOR (combustion) has been implemented. These results are consistent with the EOR experiences at Target 1 described in Section 3.1. Our weighted-distance criterion ranks as most similar cases (a) the Pembina-Nisku “O” pool, and (b) the Rainbow-Keg river “F” pool, both of which are light-oil reservoir located in Alberta, Canada. As reported by Galas et al. (2012 and references therein), miscible gas flooding at (a) started in 1986 through one vertical injector and was completed in the early nineties, when the production started to decline. The incremental recovery factor due to gas injection was estimated as equal to 40% of the Original Oil In Place (OOIP) in this pool. Tertiary HC miscible flood at (b) started in 1996 in the NW lobe and in 2000 on the whole pool. This EOR production scheme is still active and economically profitable, extending over 12 injectors and 39 production wells. A total of 38% of the OOIP, was produced from primary recovery, whereas the incremental recovery factor due to HC flooding is currently estimated as providing 15% of the OOIP (Galas et al., 2012). The Combustion project amongst the nearest neighbors of Target 1 is associated with Pennel field, in Williston Basin, US. This field is characterized by very light oil. As documented in Sheng (2013), the main factors leading to a successful in-situ combustion project in light oil reservoirs are (i) high reservoir temperature, (ii) high dip and (iii) relatively high oil saturation. All these requirements are fulfilled at Williston Basin, where several combustion projects have been implemented in fields providing very low primary recovery factors.

Target 3.

The results obtained for this blind test highlight the importance of combining clustering with distance-based ranking for an effective screening. Even as the information obtained from cluster analysis would favor the use of a miscible gas injection as the most suited technique for Target 3, the evaluation of weighted distances within the cluster leads to a different conclusion. As indicated by the analog ranking depicted in Fig. 12c, the whole sub-set of chemical (mostly polymer) EOR projects belonging to Cluster 2 are concentrated in the nearest neighborhood of Target 3. This conclusion is consistent with the actual EOR technique that has been successfully implemented at the site (Section 3.1). Amongst these polymer projects, the reservoir which is nearest to the target is the Shuanghe field, in China. This sandstone reservoir had a long history of water flooding when the first polymer injection pilot was performed, in the early 90s. Chemical EOR has been practiced in China since the mid-80s (e.g., in the Daqing field). However, it has to be noted that the Shuanghe pilot required an ad-hoc design, because of the high reservoir temperature (about 160 °F), as compared with field conditions characterizing prior polymer flooding tests. A new formula of Hydrolyzed polyacrylamide (HPAM) developed to enhance thermal stability of the polymer was employed at Shuanghe for the first time, resulting in a 10% increase in oil recovery (He et al., 1998). The field is currently treated with Surfactant-polymer flooding. First tests of a new high-temperature resistant foam with low interfacial tension provide promising results in improving post-chemical-flood recovery (Tang et al., 2014). Reservoir temperature is a critical property for chemical EOR applications. As shown by Table 2, temperature at Target 3 is estimated at 170 °F. Hence, the experience at Shuanghe can provide useful information for further development of the EOR project at the Target. It can be noted that a hydrocarbon Miscible WAG (MWAG) injection project, i.e., the Kuparuk river field, in Alaska (US), is ranked as the highest analog to Target 3. Miscible injection at Kuparuk river started in 1988 and was expanded to a large scale since 1996, becoming one of the largest MWAG projects worldwide (Jensen et al., 2012). Although miscible gas injection is the primary EOR approach implemented in this field, chemical EOR methods also constitute a potentially viable alternative: laboratory tests and field trials have suggested low-salinity water injection to be a promising EOR technique within all of the Alaskan North Slope fields (McGuire et al., 2005, Llano et al., 2013).

4. Summary and conclusions

We develop and test a new screening method conducive to the identification of the most suitable EOR technology on target reservoirs. Our method relies on the following key components:

1. A database of thermal (in-situ-combustion, hot-water and steam-injection projects), chemical (polymer, surfactant or Alkali-Surfactant-Polymer flooding) and gas/WAG injection (miscible/immiscible; Hydrocarbon, Nitrogen, Carbon dioxide) EOR projects operated worldwide and collected from an extensive review;

2. A Principal Component Analysis applied to data and employed not only for dimensionality reduction but also for an accurate choice of distance metrics to define similarity between projects;

3. A two-step algorithm to assess analogy between data and targets, on the basis of oil and reservoir properties.

A blind test of the approach has been performed by considering three field cases where given EOR techniques have been selected. The degree of documented success demonstrates the effectiveness of our methodology to capture the optimal EOR techniques for the selected fields. We note that the results of the blind test are also consistent with the application of conventional screening criteria. This can be seen by mapping our three target fields in the space of reservoir depth versus oil viscosity (see Table 1 and Fig. 1). The objective of the advanced screening method we propose is not limited to provide a qualitative identification of the most suitable EOR technique for a given target field. A major strength of our method is that it is conducive to a quantitative assessment of the analogy between target and database fields, where EOR techniques have already been implemented. This quantification cannot be achieved by conventional screening approaches. In this context, identification of analogs is key to the decision making process. The results obtained for Target 3 constitute a stark example of the importance of analogs evaluation and highlight also the key importance of our distance-based ranking, as applied subsequently to cluster decomposition, to refine the search for analogs.

The complete collection of documentation related to worldwide EOR experiences is included in our database, thus enabling one to acquire from the earliest stages of the project evaluation a technical know-how on the most promising EOR technique.

References

Abbas, E. and Song, C.L. 2001. Artificial Intelligence Selection with Capability of Editing a New Parameter for EOR Screening criteria. Journal of Engineering Science and Technology 6 (5): 628-638.

Al Adasani, A. and Bai, B. 2011. Analysis of EOR projects and updated screening criteria. Journal of Petroleum Science and Engineering 79 (1-2): 10-24.

Alvarado, V. et al. 2002. Selection of EOR/IOR Opportunities Based on Machine Learning. Paper SPE 78332 presented at the SPE 13th European Petroleum Conference, Aberdeen, 29–31 October 2002.

Alvarado, V. and Manrique, E. 2010. Enhanced oil recovery: field planning and development strategies. Elsevier Inc., Oxford.

Anikin, I. 2014. Knowledge Representation Model and Decision Support System for Enhanced Oil Recovery Methods. Paper presented at the International conference on Intelligent Systems, Data Mining and Information Technology, Bangkok, 21-22 April.

Babushkina, E.V., Rusakov, V.S., Rusakov, S.V., Shchipanov, A., Berenblyum, R.A. and Khrulenko, A.A. 2013. Forecasting IOR/EOR Potential Based on Reservoir Parameters. Paper P22 presented at the 17th European Symposium on Improved Oil Recovery, St. Petersburg, 16-18 April.

Daszykowski M., Kaczmarek, K., Vander Heyden, Y. and Walczak, B. 2007. Robust statistics in data analysis – A review Basic concepts. Chemometrics and intelligent laboratory systems 85 (2): 203-219.

Galas, C.M.F., Clements, A., Jaafar, E., Jeje, O., Holst, D. and Holst, R. 2012. Identification of enhanced oil recovery potential in Alberta, Phase 2. Final report for ERCB.

Gharbi, B.C.R. 2000. An expert system for selecting and designing EOR processes. Journal of Petroleum Science and Engineering 27 (1-2): 33-47.

He, J. et al. 1998. High Temperature Polymer Flooding in Thick Reservoir in Shuanghe Oilfield. Paper SPE 50933 presented at the SPE International Oil and Gas Conference and Exhibition Beijing, 2-6 November.

Heller K. and Ghahramani, Z. 2005. Bayesian hierarchical clustering. Paper presented at the ICML ’05, 22nd International Conference on Machine Learning, Bonn, 7-11 August.

Kamari, A., Nikookar, M., Sahranavard, L. and Mohammadi, A.H. 2014. Efficient screening of enhanced oil recovery methods and predictive economic analysis. Neural Computing and Application 25 (3-4): 815-824.

Koottungal, L. 2008. 2008 Worldwide EOR Survey. Oil & Gas Journal 106 (4): 44–59.

Koottungal, L. 2010. 2010 Worldwide EOR Survey. Oil & Gas Journal 108 (14): 36-53.

Koottungal, L. 2012. 2012 Worldwide EOR Survey. Oil & Gas Journal 110 (15): 47–59.

Koottungal, L. 2014. 2014 Worldwide EOR Survey. Oil & Gas Journal 112 (5): 79–91.

Jensen, T.B. et al. 2012. Kuparuk River Unit Field- The First 30 Years. Paper SPE 160127 presented at the SPE Annual Technical Conference and Exhibition, San Antonio, 8-10 October.

Llano, V. et al. 2013. Water Management for EOR Applications - Sourcing, Treating, Reuse and Recycle. Paper OCT 24199 presented at the Offshore Technology Conference, Houston, 6-9 May.

Manrique, E., Izadi, M., Kitchen, C.D. and Alvarado, V. 2009. Effective EOR Decision Strategies With Limited Data: Field Cases Demonstration. SPE Reservoir Evaluation & Engineering 12 (4): 551-561.

Manrique, E. et al. 2003. Perspectives of CO2 injection in Venezuela. Paper presented at the 24th Annual Workshop and Symposium for the IEA Collaborative Project on Enhanced Oil Recovery, Canada, 7–10 September 2003.

McGuire, P.L. et al. 2005. Low Salinity Oil Recovery: An Exciting New EOR Opportunity for Alaska's North Slope. Paper SPE 93903 presented at the SPE Western Regional Meeting, Irvine, 30 March-1 April.

Moreno, J., Gurpinar, O. and Liu, Y. 2014. EOR Advisor System: A Comprehensive Approach to EOR Selection. Paper IPTC 17798 presented at the International Petroleum Technology Conference, Kuala Lumpur, 10-12 December.

Sheng, G. 2013. Enhanced Oil Recovery Field Case Studies, first edition, 712. Elsevier Inc., Oxford.

Sirinukunwattana, K., Savage, R.S., Bari, M.F., Snead, D.R.J. and Rajpoot, N.M. 2013. Bayesian Hierarchical Clustering for Studying Cancer Gene Expression Data with Unknown Statistics. PLoS ONE 8 (10): e75748.

Surguchev, L.M. and Li, L. 2000. IOR Evaluation and Applicability Screening Using Artificial Neural Networks. Paper SPE 59308 presented at the SPE/ DOE Improved Oil Recovery Symposium, Tulsa, 3 – 5 April.

Taber, J.J., Martin, F.D. and Seright, R.S. 1997 (a). EOR Screening criteria revisited – Part 1: Introduction to Screening Criteria and Enhanced Recovery Field Project. SPE Reservoir Engineering 12 (3): 189-198.

Taber, J.J., Martin, F.D. and Seright, R.S. 1997 (b). EOR Screening criteria revisited – Part 2: Implications and impact of Oil Prices, SPE Reservoir Engineering 12 (3): 199-206.

Tang, S. et al. 2014. A Novel Low Tension Foam Flooding for Improving Post-Chemical-Flood in Shuanghe Oilfield. Paper SPE 169074 presented at the SPE Improved Oil Recovery Symposium, Tulsa, 12-16 April.

Trujillo, M. et al. 2010. Selection Methodology for Screening Evaluation of Enhanced-Oil-Recovery Methods. Paper SPE 139222 presented at the Latin American & Caribbean Petroleum Engineering Conference, Lima, 1-3 December.

Uribe, J.P. et al. 2010. Innovative methodology to revitalize a heavy-oil mature field by identifying opportunities to apply new cycles of steam injection. Paper SPE 138918 presented at the SPE Latin American and Caribbean Petroleum Engineering Conference, Lima, 1-3 December.

Vega Riveros, G. and Barrios, H. 2011. Steam injection experiences in heavy and extra-heavy oil fields, Venezuela. Paper SPE 150283 presented at the SPE Heavy Oil Conference and Exhibition, Kuwait City, 12-14 December.

Yuqiu, L. and Yali, Z. 2009. Reservoir Screening Criteria for Heavy Oil Thermal Recovery in Liaohe Oilfield. China Oil & Gas 2: 31-35.

Zerafat, M.M., Ayatollahi, Sh., Mehranbod, N. and Barzegari, D. 2011. Bayesian Network Analysis as a Tool for Efficient EOR Screening. Paper SPE 143282 presented at the SPE Enhanced Oil Recovery Conference, Kuala Lumpur, 19–21 July.

Tables and Figures

	ϕ (%)	k (mD)	D (ft)	API (°)	μ (cp)	T (°F)
Min	3	0.1	200	8	0.1	45
Max	40	11500	16150	57	5×10^{6}	290

Table 1. Ranges of oil and reservoir properties characterizing the database elements.

	ϕ (%)	k (mD)	D (ft)	API (°)	μ (cp)	T (°F)
Target 1	14	200	8760	42	0.23	185
Target 2	29	3000	1300	8.5	2500	118
Target 3	22.5	100	5250	16	23	170

Table 2. Oil and reservoir properties characterizing the test cases.

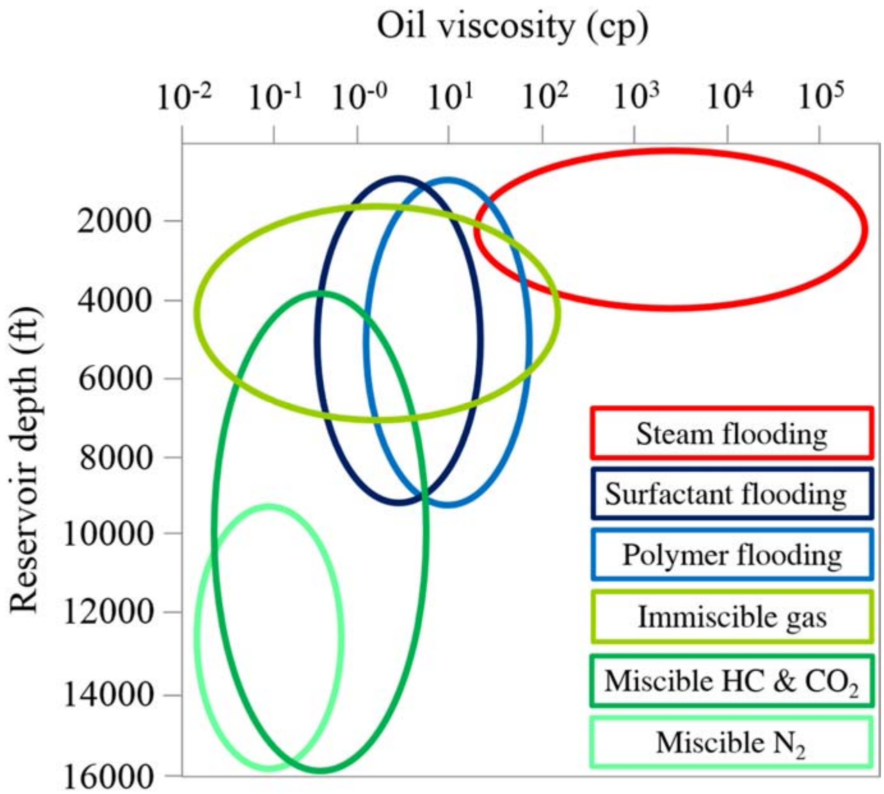

Figure 1. Conventional screening criteria inferred by Taber et al. (1997a,b): colored curves demarcate the regions of applicability of the main EOR techniques in the phase space of reservoir depth versus oil viscosity.

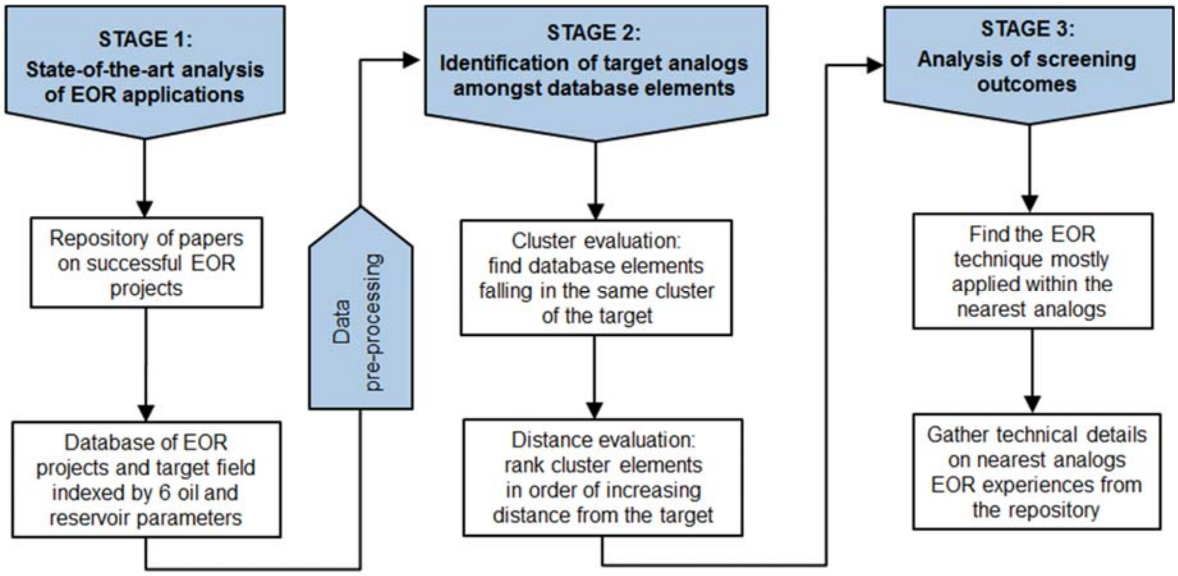

Figure 2. Workflow of the EOR screening procedure.

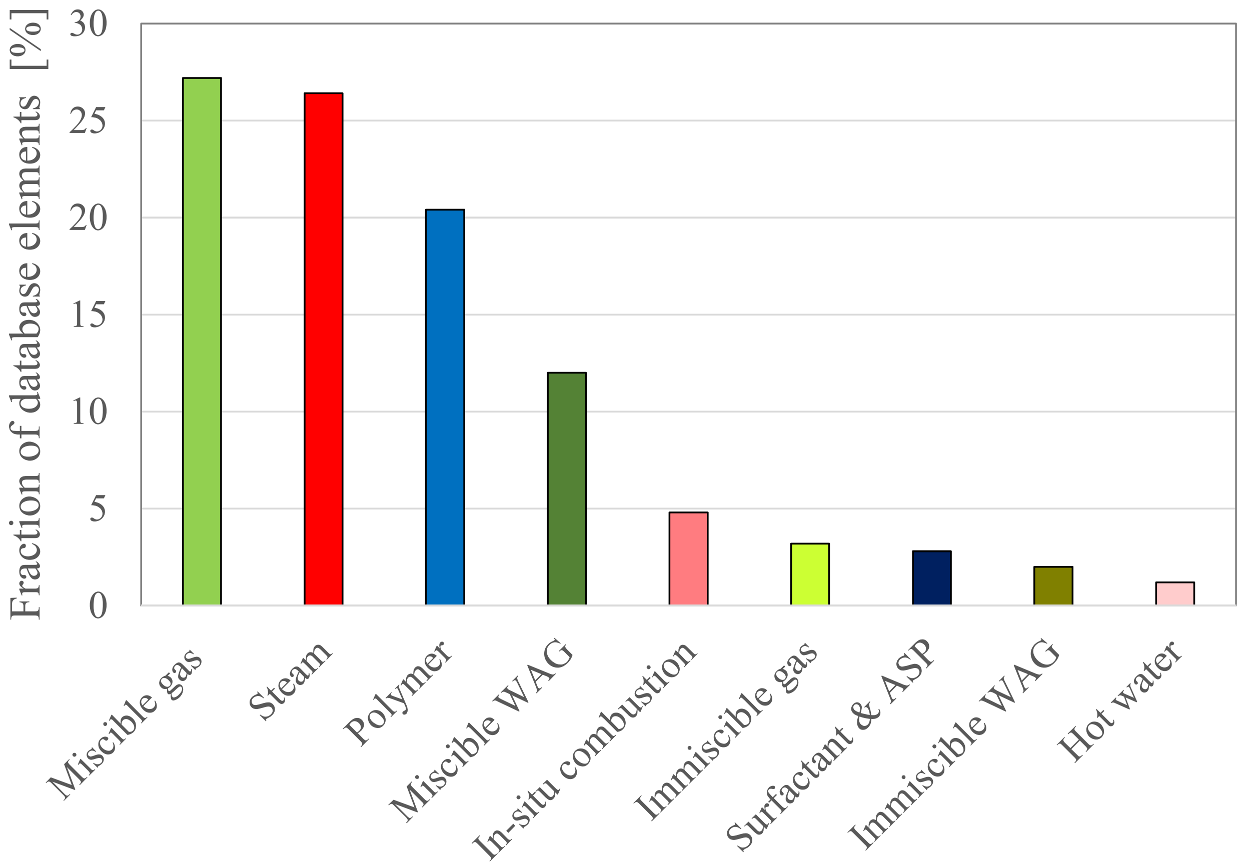

Figure 3. Relative occurrence of EOR techniques across the 250 samples of the database.

Figure 4. Database of EOR projects projected onto a world map. Color scale represents the associated EOR technique.

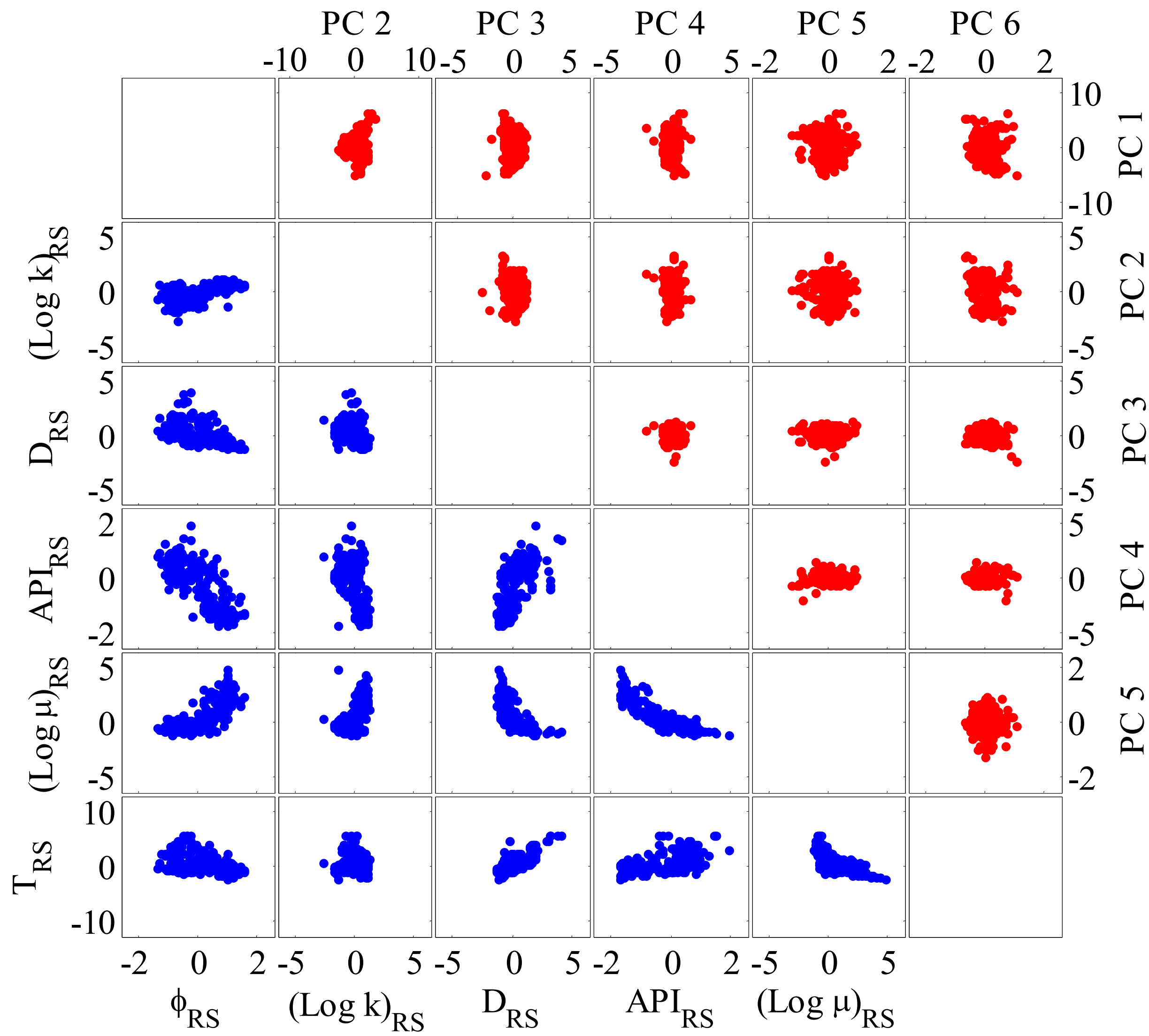

Figure 5. Scatterplots of robust-standardized properties (below diagonal) and Principal Components (above diagonal).

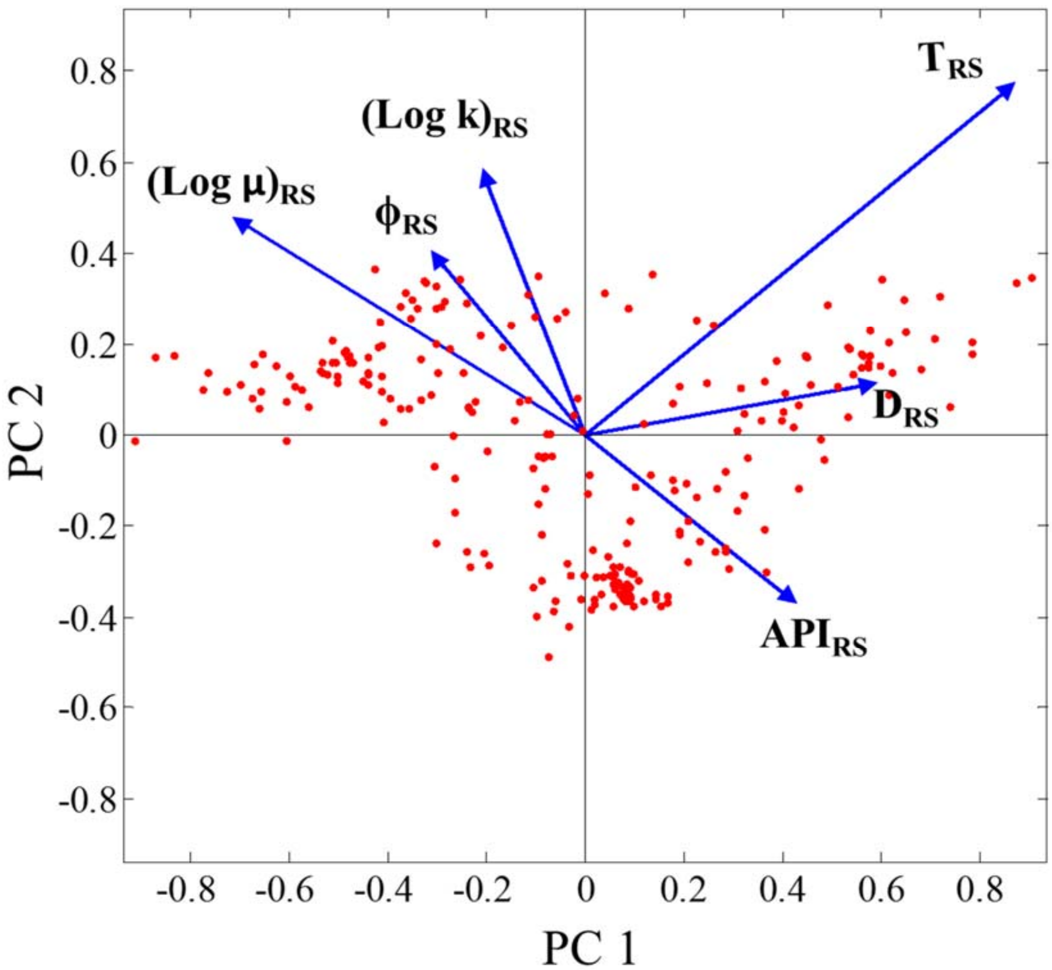

Figure 6. Effect of Principal Component Analysis: blue arrows are vectors indicating the contribution of each of the RS-variables to the first two PCs. Red symbols represent database elements in this space.

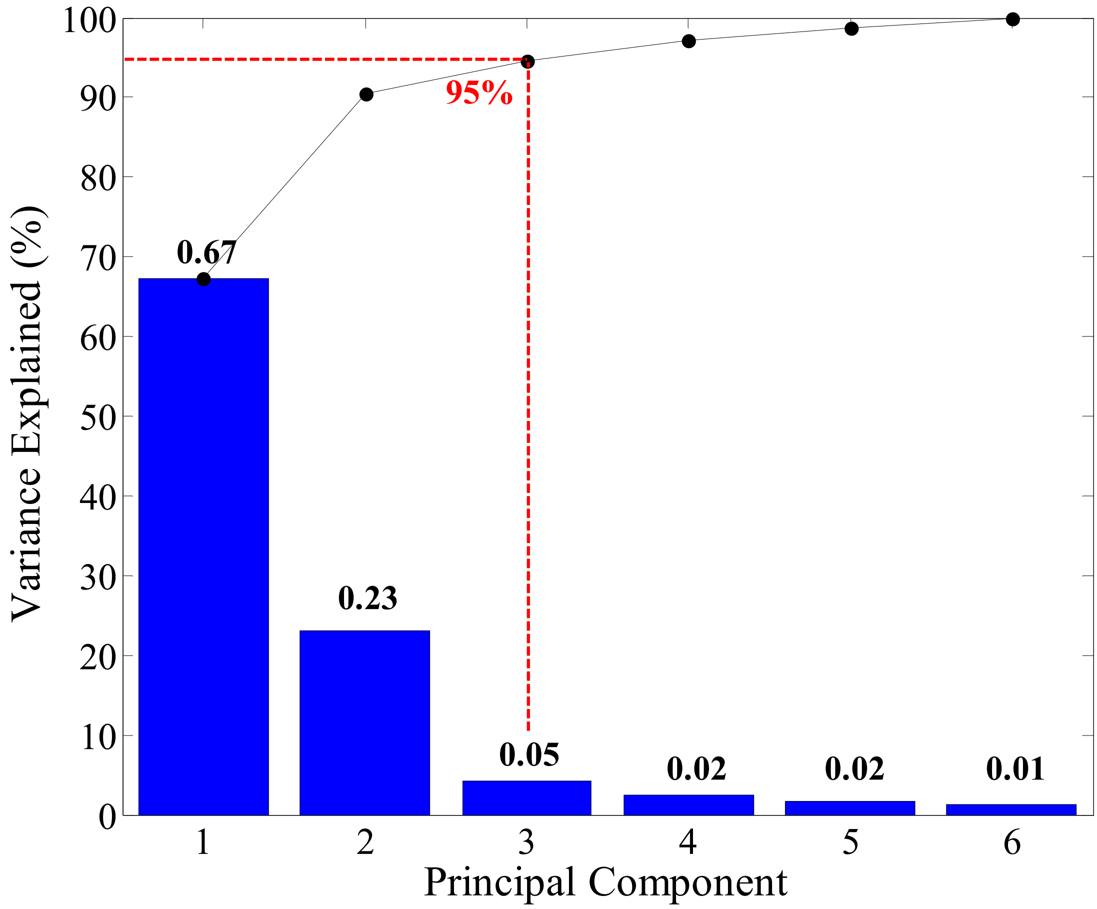

Figure 7. Fraction of variance explained by each PC. Space dimensionality can be reduced from 6 to 3 with no appreciable loss of information.

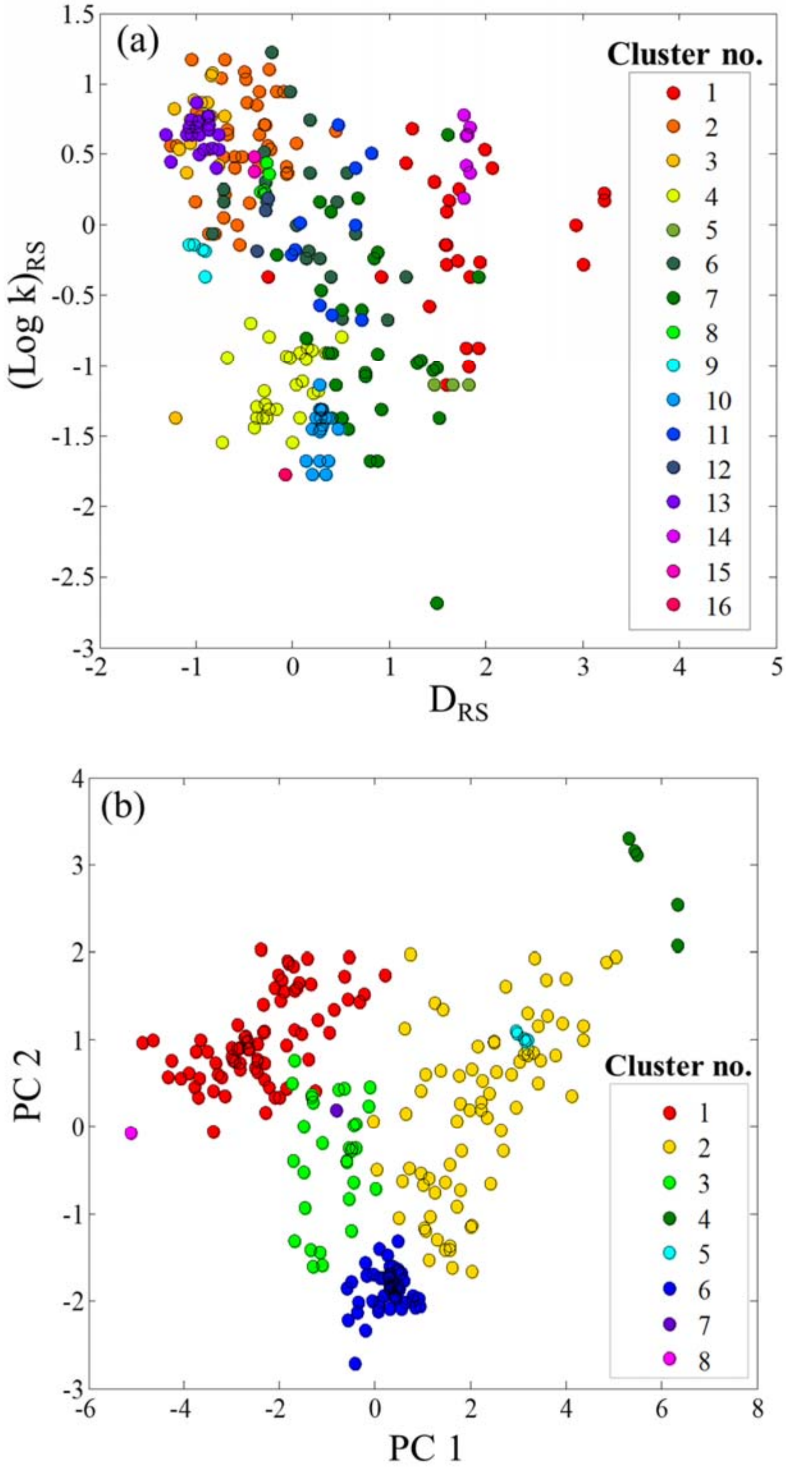

Figure 8. Results of the BHC algorithm applied to (a) the space of the Robust-standardized variables and plotted in the phase space (Log k)RS - DRS; (b) the reduced space of PCs.

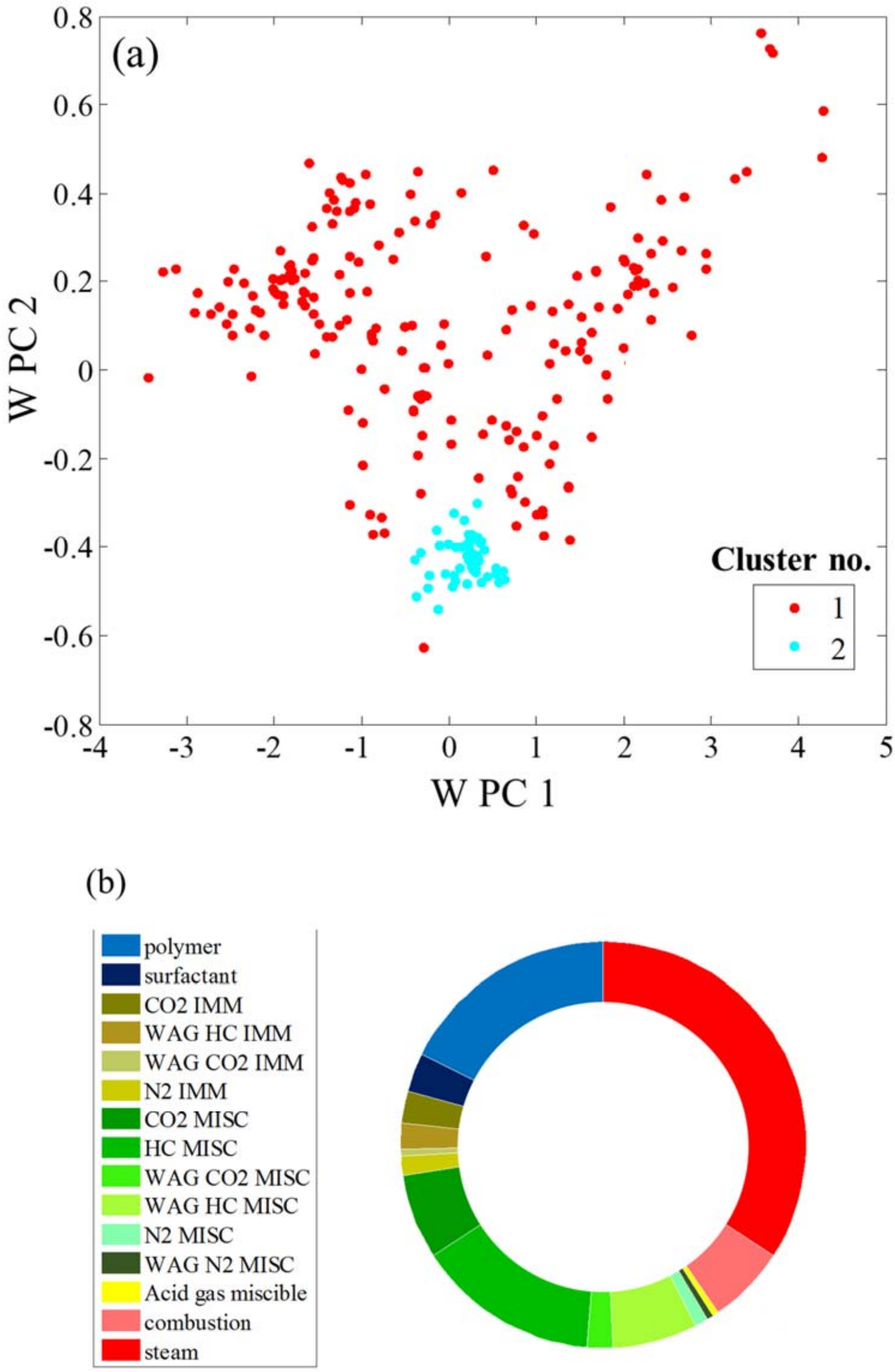

Figure 9. (a) BHC algorithm applied to the space of the PCA-weighted principal components W PC1 and W PC2; (b) Proportion of EOR techniques associated with elements in Cluster 1.

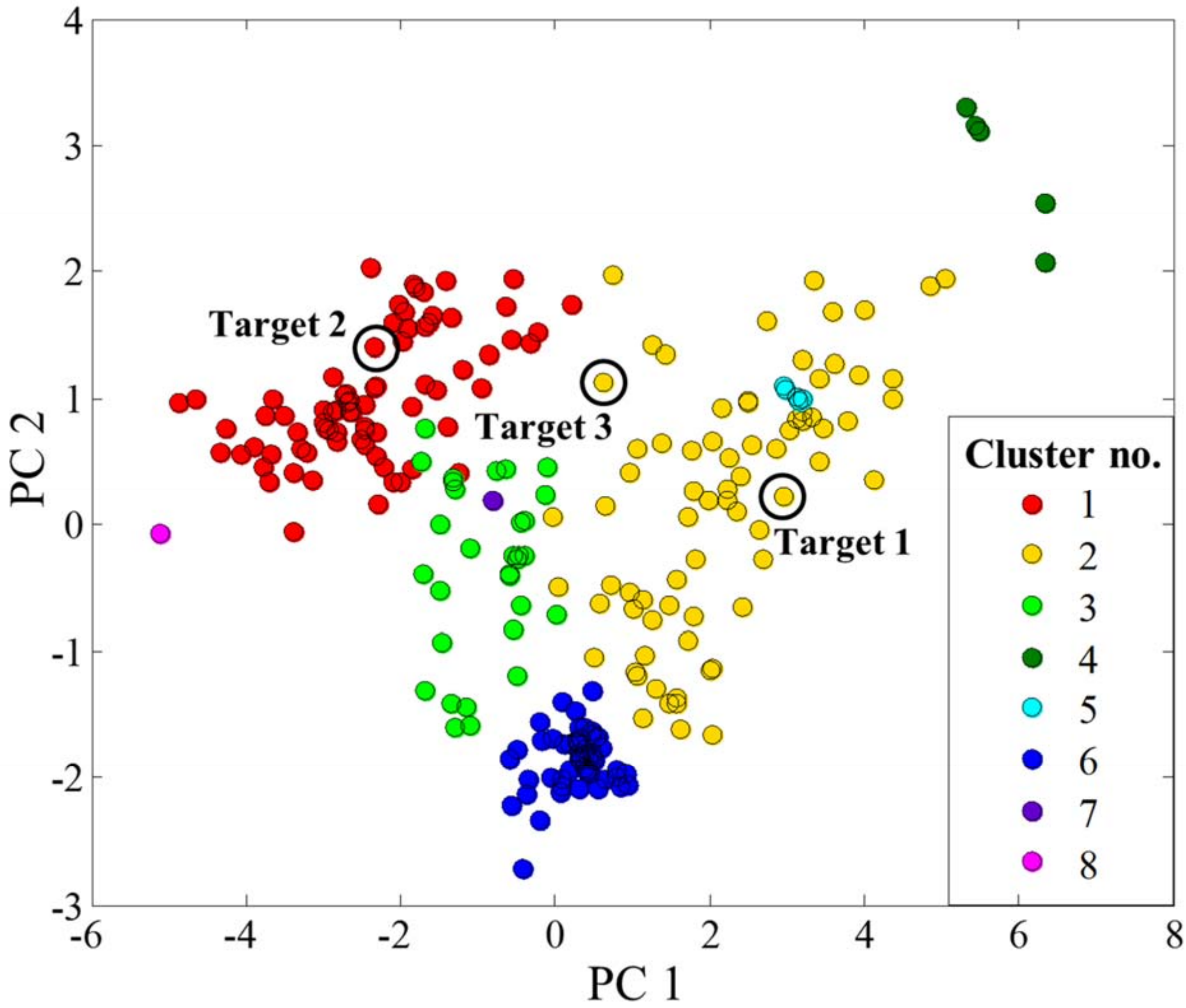

Figure 10. Scatterplot of database elements and targets in the space of the first two PCs. Dots color identifies the 8 clusters obtained from the BHC algorithm.

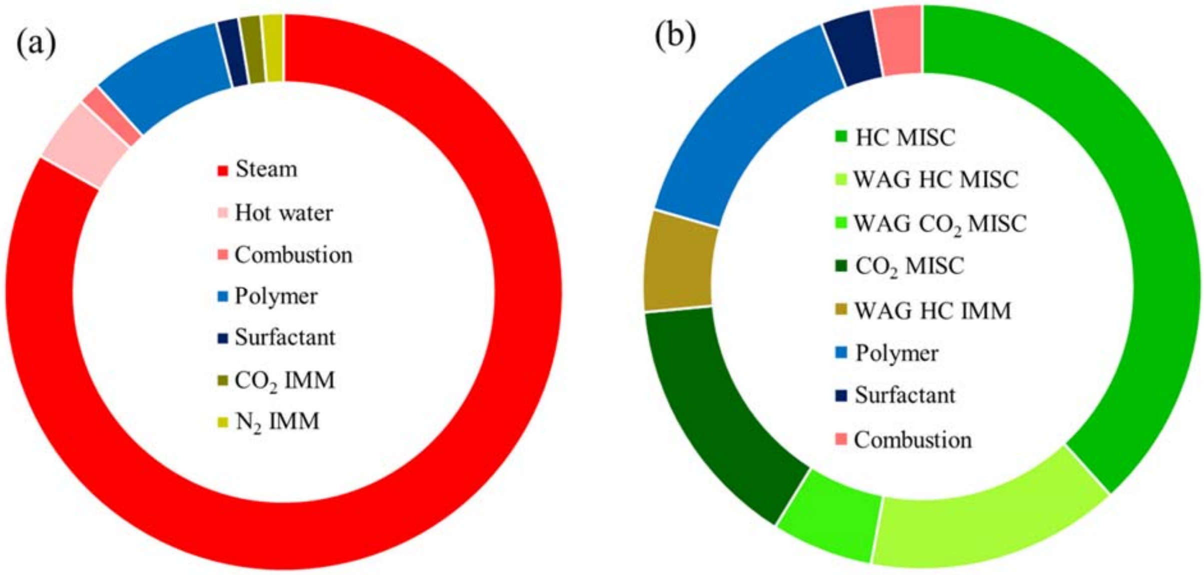

Figure 11. Screening results: proportion of EOR techniques associated with elements in (a) Cluster 1 (i.e., Target 2 analogs) and (b) Cluster 2 (i.e., Targets 1 and 3 analogs).

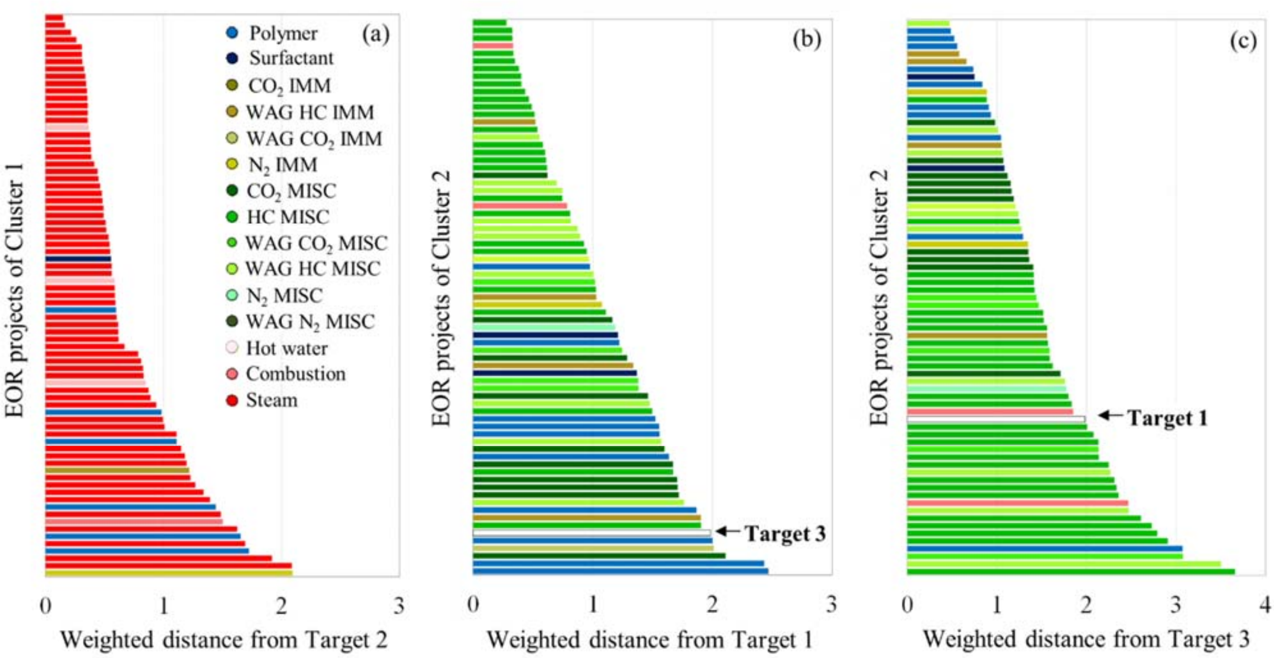

Figure 12. Screening results: distance-based ranking of cluster elements and associated EOR techniques for (a) Target 2 (i.e., Cluster 1 elements), (b) Target 1 (i.e., Cluster 2 elements), and (c) Target 3 (i.e., Cluster 2 elements).