ABSTRACT

New challenges are facing the petroleum industry: processing heavier crude oils with more sulfur; needs for better environmentally compatible utilization of fractions and residuals into value added products; formulation of fuels and lubricants with more expensive alternative components; tougher competition; appearance of hybrid and fuel cell vehicles, which require different formulation strategies; tougher environmental restrictions in the “upstream” and “downstream” activities of petroleum companies, etc.

The main principles of chemical product design for fuels and lubricants, which are additive products, built up from base stocks, components and additives are presented in this review.

The presently used methods of the petroleum industry for process and product design, including those embedded in process simulators, are critically analyzed and relevant problems are pointed out. The new methods for design of properties of molecules and mixtures (QSPRs, process design of discrete and continuous systems, “properties – property” correlations for screening, etc.) being developed in the leading laboratories of the private and the public sector, and their relevance to the new challenges are reviewed.

The review also identifies the problems that face the implementation of the new methods and discusses opportunities for solving them.

Keywords: fuels, lubricants, additives, process and product design, QSPR.

INTRODUCTION

Fuels and lubricants are important components of modern civilization. In the beginning of the 21st century, EU and the rest of the industrialized nations adopted strategies for solving simultaneously both the energy and environmental challenges which the world is facing [1, 2]. The respective EU directive, popular as directive “20/20/20”, targets the achievement of 20% renewable energy, 20% reduction of the emissions of green-house gases and 20% improvement of energy efficiency in the year 2020, as compared to 1990 [3].

The European transport sector is 98% dependent on fossil fuels and accounts for around a quarter of the EU emissions of CO₂, Transportation and the other activities in which fossil fuels are involved (energy production, industry, etc.), contribute most of the rest of the criteria and specific air pollutants – SO₂, NOₓ, PAH, particulate matter, etc., so the evaluation of the impact of any energy/environment improvement option should be done on the basis of its life cycle analysis (LCA) [4, 5]. LCA requires that the impact of a transportation fuel, for instance, should be estimated not only on the basis of the pollutants emitted from the vehicle, but should start with the pollution involved in the production and transportation of the raw materials (e.g. crude oil or biomass), from which it is produced, its own production, transportation, storage, marketing, usage and so on. Thus, the petroleum industry is always involved.

Biofuels are the key to the EU directive requiring 10% of all transportation fuels to be renewable by 2020, and the vision of achieving 25% in 2030 and
beyond, without compromising agricultural production [6]. Those used today include bioalcohols, esters of fatty acids (biodiesel) and biogas. Biohydrogen for fuel cells and “biomass to liquids” synthetic fuels are targets for future development [7].

Lubricants are directly related to improving energy efficiency, but they also pollute the environment. Motor oils, for instance, are partially burned in engines and thus influence both the composition of emitted pollutants, and the efficiency of after-burning catalysts. They can also leak, e.g. from agricultural equipment or marine vessels, so biodegradable oils are also being developed.

The above tasks pose new challenges to modern refineries [8-10], some of which include:

- Heavier crude oils with more sulfur have to be processed to products with tougher environmental restrictions (e.g., less than 10 ppm sulfur and less than 11 % polyarenes in ultra low sulfur diesel fuels).
- Higher crude oil prices require better environmentally compatible utilization of fractions and residuals into value added products, in order to achieve reasonable profit margins.
- Formulation of fuels and lubricants with alternative (e.g., “gas to liquid, GTL”, “coal to liquid, CTL”) and renewable synthetic components (e.g., “biomass to liquid, BTL”), which are more expensive and often produced outside of the refinery, not only puts up prices, but might also require new relationships for optimization of the final composition.
- Tough competition is a fact not only in the additive market, but also in the processing market, traditionally dominated for many years by the leading international companies.
- Existing challenges from producers of renewable fuels and emerging challenges related to hybrid and fuel cell vehicles can drastically change the competition in the fuel and lubricants market.
- Tougher environmental restrictions on pollution on every stage in the “upstream” and “downstream” activities of petroleum companies are being imposed.

In order to meet the above challenges the petroleum companies need wider implementation of information-based computer aided technologies, which can optimize on-line both production processes (“process design”) and quality of products (“product design”). The availability of such technologies depends mainly on the availability of methods and devices for on-line analysis of production streams and control of pre-set parameters, and algorithms for process and product design for programming the devices. The development of the above components is in the mainstream of chemical engineering and involves also education in process and product design, which is part of the curricula of many universities [11-14].

Chemical process and product design have been recently defined as the two main interrelated parts of chemical product engineering [15]. Obviously, discussing all aspects of chemical product engineering is an immense and product specific task. The aim of this review article is limited mainly to the analysis of existing and the development of new computational relationships, relevant to fuels and lubricants design. Within this focus, the contribution of the University of Chemical Technology and Metallurgy will be also outlined.

PETROCHEMICALS AND ADDITIVE PRODUCTS

A refinery offers to the market mainly two types of products – petrochemicals and additive products. Petrochemicals are pure chemicals like benzene, toluene, phenol, acetone, etc. and relatively pure mixtures of structurally similar compounds like glycols, amines and ethanolamines, polymers, etc.

Marketed fuels and lubricants are additive products, which consist of a base (fuel, oil, etc.), components (bio-, synthetic, etc.), and synthesized additives.

The base fuels and oils are typically blends of crude oil fractions (base stocks), produced by different refining processes [16, 17]. However, the base stocks of fully synthetic base oils are entirely synthetic mixtures.

The components are usually synthetic products (e.g., alkylates, ethers, GTL, BTL, polyalphaolefins, esters, etc.) synthesized from lower molar mass crude oil hydrocarbons, biomass and natural esters (e.g., biodiesel, bioalcohols, bioesters, etc.) or coal liquids [18-21].

The additives are relatively pure synthesized organic compounds, which contain heteroelements (i.e., different from C and H) – O, N, S, P, metals, etc. Many additives contain more than one heteroatom in their molecules [22, 23]. They may influence more than one functional property of the fuel or lubricant, which does not necessarily depend on the type and amount of heteroelements they contain. This is important in formulation because the relevant requirements of a speci-
fication might be achieved, by additives with significantly different chemical structures.

The most widely used additives are either soluble or form stable colloid solutions in the base stock. However, solid additives insoluble in the base (nanopowders, “soft” metals, etc.), which form stable dispersions, are becoming increasingly popular as friction modifiers for greases and oils [21-24].

Base stocks, components and additives, for the purposes of this review, would be further distinguished by their influence on the properties of the finished product.

Base stocks are the largest amount of the finished product and determine mainly its bulk properties (e.g., density, viscosity, etc.). They are prepared by blending particular production streams, according to base stock specifications. Components are used in smaller amounts than base stocks, but their concentrations are considerably larger than, those of the individual additives. Their effect on the properties of the finished product is proportional to their concentration and they influence the bulk properties of the finished product by linear or nonlinear interaction with base stocks. For instance, the use of oxygenates at an allowed total concentration of up to 15 % in motor gasoline, requires reformulation of its hydrocarbon composition in order to achieve the specified vapor pressure, distillation characteristics, etc.

The functional effect of an additive on the performance of the finished product typically passes through a maximum, which is realized at relatively low concentrations and it does not influence the product’s bulk properties. However, the total amount of the additives (the so called “package”), needed to achieve a targeted specification of a motor fuel, and especially – of a targeted performance grade of a motor or gear oil, can be comparable to the amount of a synthetic component used in the finished product [25].

The above definition of additive products can be applied not only to fuels and lubricants but can be adapted to other finished products of the industrial chemical technology, based on polymers or elastomers, fertilizers, pesticides, drugs, etc. [26].

**DESIGN OF PROCESSES FOR PRODUCTION OF BASE STOCKS, COMPONENTS AND ADDITIVES**

As explained above, the design of fuels and lubricants involves design of the processes for the production of the base stocks, the components and the additives.

**Production of base stocks by refining of crude oil fractions**

Fuel and lubricant base stocks are produced from crude oil by processes that can be classified as processes for physical separation and – processes for chemical conversion [27, 28] (Table 1).

Process design today is done with process simulators, which have become an essential tool also for performance evaluation of existing processes and equipment. There are excellent published texts devoted to process design, which describe in detail how process simulators are developed, the theory and algorithms they are based upon, examples for their application in chemical engineering, etc. [11, 13, 29-31]. Modern refining technology books give also examples of the application of process simulators in refineries [28, 32]. The primary task of a process simulator is to serve as an efficient tool for determining the best technological parameters and installations for transforming given raw materials into products with specified properties. Within this context, the specified properties of the products of a particular refining installation might be only distantly

---

**Table 1. Examples of modern processes for production of base stocks from crude oil.**

<table>
<thead>
<tr>
<th>Physical separation</th>
<th>Chemical conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distillation</td>
<td>Thermo-catalytic</td>
</tr>
<tr>
<td>Solvent extraction</td>
<td>Catalytic reforming</td>
</tr>
<tr>
<td>Solvent dewaxing</td>
<td>Fluid catalytic cracking</td>
</tr>
<tr>
<td>Solvent deasphalting</td>
<td>Hydrocracking</td>
</tr>
<tr>
<td>Dewaxing with urea or molecular sieves.</td>
<td>Hydroisomerization</td>
</tr>
<tr>
<td></td>
<td>Catalytic dewaxing</td>
</tr>
</tbody>
</table>

219
related to the properties required from the finished fuel or lubricant, which goes to the market.

Within the defined aim of this presentation we would note several features of the present simulators, which concern the quality of base stocks, as ingredients of the finished product:

- The precision of modeling and simulations is determined by the uncertainties of the data and the approximations of the models they employ. Their feasibility for particular application should be estimated also within the limitations imposed by the application. For instance, TBP data with relatively low precision might be useful for on-line monitoring of a distillation column.
- Most of the process simulators have huge databases of the property data for individual compounds, needed for design. Some of these are experimental, but care should be taken when using them, because errors even in the best databases have been reported [33, 34]. Clear distinction between experimental and predicted data in databases is important, but not always available.
  - If experimental data are not available, they are predicted by choices of different methods, as provided by the simulator. For pure compounds these relationships are usually based on empirical correlations, describing molecular structures by group contribution methods. Their limitations will be discussed later.
  - Crude oils and their fractions are typically presented by up to 50 pseudo-components and a smaller number of chemically identifiable lumps, obtained from crude oil assays, as described, for instance, in [28]. In the light of the first feature, outlined above, for most applications this is adequate for the present successful use of process design.
  - Pseudo-components of petroleum fractions, defined by dividing the boiling temperature curve, model adequately the processes of physical separation. For processes with chemical transformation of the feed, chemical identification by lumps (groups of compounds with similar molecular structure) is required in order to model their chemical transformations [28, 32, 35].
  - The thermophysical properties, needed for design (e.g., pseudocritical properties, etc.) with crude oil fractions, are predicted with methods, embedded in the process simulators. These methods use semi-empirical and empirical “property – property” relationships, derived by fitting properties of fractions from a limited number of crudes. The same holds for prediction of properties, which are relevant to the specifications of the finished fuels and oils – octane numbers and Reid vapor pressure for gasoline base stocks, cetane numbers for diesel fuel base stocks, viscosity of lubricant base stocks, etc. Their limitations will be discussed later.
  - The development of new methods and devices for analyses of process streams creates new opportunities. Presently used monitoring devices employ gas chromatography (GC) and Fourier transformation infrared spectrophotometry (FTIR) [36]. The possibility for future use of these and other methods in in-line and stand-alone devices in fuels and lubricants product design will be also discussed later.

Design of processes for synthesis of components

As already noted, finished fuels and lubricants contain synthetic hydrocarbon and non-hydrocarbon components [19, 21, 25, 37]. Typical examples for gasolines are the high octane number components: alkylates, oligomerizates, C5 – C6 isomerizates, ethers, low-molecular mass alcohols, GTL and BTL Fischer-Tropsch (F-T) products, etc. Diesel fuels presently employ less synthetic components (e.g., GTL), but chemical- and biotechnologies for co-processing of crude oil fractions together with BTL and CTL are being developed [21, 38, 39]. Biodiesel, when used in up to 5% concentration in Europe, acts as an additive (especially for its lubricating effect) rather than as a component. However, increasing its concentration to 20% (e.g., in the US) leads to driving of the blended fuel properties out of specification. The same holds for the amount of alcohols in gasoline.

Synthetic lubricant components and base stocks can be hydrocarbon (e.g., polyalpha- and polyinternal olefins, F-T products, etc.) or non-hydrocarbon compounds – modified natural esters, esters of polyvalent alcohols with fatty acids, etc. [19]. Fig. 1 shows generalized structures of diesters, used as synthetic lubricant base stocks and/or components.

Syntheses of hydrocarbon gasoline octane boosters, ethers (MTBE and others), etc. are typically carried out in the refineries and the computer simulators can be expanded with the respective modules for modeling isomerization, alkylation, and even – the production of ethers with reactive distillation.
Fig. 1. Typical structures of synthetic esters used in lubricants.

Depending on the capabilities of its chemical complex a larger refinery might produce some non-hydrocarbon fuel and/or lubricant components, and starting molecules for their synthesis. Regardless of the place of production, the modules for modeling the processes in this case, are usually not available in software products for specialty chemicals [26], and relationships for prediction of syntheses relevant properties still have to be developed [40 - 43].

We would like to note two common features of the design of synthetic components, which are important for this presentation. Firstly, the chemical structures of the starting and the produced molecules, unlike those in crude oil fractions, are chemically defined, so “structure – property” relationships at molecular level, which are going to be discussed later, can be used. Moreover, the composition of the components can be more easily monitored by GC, FTIR, chemical analyses and other methods. Secondly, the performance properties contribution by the synthetic components in the finished products can be more easily modeled. For instance, the composition of gasoline components can be analyzed fully in-line by GC, and related to predicted values of octane number, vapor pressure, etc. The same holds for the yet stand-alone devices for analyses of biodiesel and most of the lubricant components.

**Design of processes for syntheses of additives**

Modern fuel and lubricant additives span a wide nomenclature of chemical compounds [23, 25, 39, 44]. Table 2 shows typical classes of additives and their main functions. Examples of chemical compounds and substances, used as additives for fuels and lubricants are presented hereunder [23-25, 39, 44]:

*Metalorganic compounds*: Zn-alkylaryl dithiophosphates, Zn alkyl dithiophosphates, Modithiophoshasites; Ba, Ca-, Mg-, Na-, long-chain alkaryl sulphonates, phenates, sulphurized phenates, alkyl salicylates, phosphonates, etc.

<table>
<thead>
<tr>
<th>Additive type</th>
<th>Primary function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersants and Detergents</td>
<td>Maintain clean surfaces by suspending and preventing from aggregation degradation, etc. products insoluble in the the fuel or oil. In fuels keep the fuel system (e.g., manifolds, injectors, valves, etc.) clean. Overbased and alkaline additives in motor oils neutralize acidic combustion products.</td>
</tr>
<tr>
<td>Oxidation inhibitors</td>
<td>React with species, produced by oxidation, to slow down chain- radical processes.</td>
</tr>
<tr>
<td>Friction modifiers, antiwear and EP agents</td>
<td>Form films improving lubrication, reducing wear or preventing welding.</td>
</tr>
<tr>
<td>Rust and corrosion inhibitors</td>
<td>Passivate metal surfaces or form a film on them. Some react with metal ions or corrosion agents.</td>
</tr>
<tr>
<td>Anti-foam and seal conditioners</td>
<td>Prevent fuels and lubricants from escaping into the environment.</td>
</tr>
<tr>
<td>Conductivity improvers</td>
<td>Disperse static electricity, accumulated in fast moving (pumping) of fuel liquids and prevent spontaneous ignition.</td>
</tr>
<tr>
<td>Combustion improvers (i.e., cetane and octane numbers, anti-smoke, NOx scavengers, etc.)</td>
<td>Improve completion of combustion of fuels and reduce concentrations of pollutants emitted to the environment.</td>
</tr>
<tr>
<td>Biocides</td>
<td>Prevent or slow down the attack on biodegradable components by microorganisms.</td>
</tr>
<tr>
<td>Pour point depressants, viscosity index improvers, etc.</td>
<td>Ensure flow at low temperatures, improve viscosity index</td>
</tr>
</tbody>
</table>
\(\text{N}^{-}\), \(\text{S}^{-}\), \text{O-containing compounds}: alkyl ammonium dithiophosphates, Alkyl dithiothiadiazoles, Alkyl and cycloalkyl nitrates and nitrites; alkaryl polyether amidoamines; alkylpolyamide amines; polyamines; polyalkyl amino phenols; polyamino polyamides, polyether amides, polyether amines, polyolefin amide alkyleneamines, polyolefin amide alkyleneamine borates, polyolefin amines, polyolefin polyamines, polyolefin succinimides, borated polyolefin succinimides, etc.

\(\text{S}^{-}\), \(\text{P}^{-}\), \text{O-containing compounds}: phosphorous sulphides: polyolefin phosphorosulphides, polyalkyl polysulphides, polyolefin sulphides, alkyl acid ester sulphides, alkenyl thioesters, etc.

\text{Polymeric Additives}: alkaryl polyethers, alkyl ester copolymers, aryl polyolefin polymers, hydrocarbon polymers, alkaryl polyethers, olefin/alkyester copolymers, polycarbonylate esters, polyalkylene glycols, polyalkyl methacrylates, polyethers, polyols, polyolefins, polyolefin esters.

\text{Dispersible compounds and substances: MoS}_2$, Teflon nanoparticles, nanodiamonds, nanopowders of low-melting metals, inorganic borates, etc.

In addition to the wide nomenclature, many of the fuel and lubricant additives are sold under particular chemical names, but are in fact mixtures of similar compounds and in many cases are diluted or solubilized in hydrocarbon or synthetic components. There is also an increasing tendency for additive companies to sell not individual compounds, but patented synergistic additive packages (mixtures), home-tuned to particular specifications. This is especially valid for lubricant additives, but also holds for some of the more sophisticated fuel additives (e.g., “clean and keep clean” formulations for engine fuel systems and injectors). Thus, the development of software programs for design of additive production processes is hindered. These features are essential for protecting intellectual property rights and the market competition [45], and would not be easy to overcome.

\text{PROPERTIES AND DATA FOR DESIGN OF FUELS AND LUBRICANTS}

In order to develop whatever relationship for predicting properties, one needs a database of measured data for the targeted properties of the targeted objects – individual compounds and/or mixtures. The objects relevant to fuel and lubricants design have already been identified as base stocks, components and additives. We may distinguish two types of desired properties – properties needed for process design and properties needed for products design.

**Property values needed for process design**

The property values for process design are parameters like normal boiling and melting temperatures, standard liquid densities, critical properties, etc. [46] Experimental data for pure compounds are collected into databases well known to chemists and chemical engineers – DIPPR, NIST, Beilstein, Dechema, etc. Estimated uncertainty should be also provided. For instance, values in the DIPPR database are supplied with maximal reliability [47]. In other databases (e.g., the NIST Chemistry Webbook) in addition to reliability, all known data for the particular property of the targeted compound are given. The databases also contain data for properties of binary and ternary mixtures, which can be used for calculation of coefficients in the so called “mixing rules”.

There are, however, properties the values of which for particular compounds cannot be measured. Typical examples are the compounds that decompose or polymerize before reaching their normal boiling temperature and/or critical state. Among the decomposing compounds are the higher molar mass fatty acids and their esters – glycerides, methyl and ethyl esters, etc., which are needed for process and product design of biodiesel or biodegradable lubricants. Compounds which polymerize include fatty acids and esters with two or three double bonds, alkenes, etc. It is not possible also to measure the liquid density at 15 °C of a compound which melts above this temperature (e.g., a long chain n-alkane).

The number of the compounds used at present by the industry, or being of its immediate interest, is estimated at around 100,000, while the chemical structures, which are theoretically possible and may eventually interest the industry in the future, are at least several tens of millions [48]. The number of the compounds for which any measured data of pure component properties are available is at most several thousands and for many properties it is much less. Moreover, values in databases are often an unidentifiable mixture of experimental data and predicted data. The reliability of predicted data provided might be unreliable, as demonstrated, for instance, by Yan et al. [49]. It should be
noted also that typically experimental data are not being measured on any systematic basis, but are rather produced ad hoc as required by a particular project.

Pure component properties are essential not only for chemical engineering computations in steady state and dynamic simulation in process and product design. They are needed also for environmental impact assessment, hazard and operability analysis, generation of species in combustion and other chain-radical processes, for creation etalon mixtures for comparison with fractions, used in GC, etc. [50 - 52].

The strategy of chemical engineering design uses the property data of the components of mixtures of pure compounds to model and simulate their behavior in production processes. The same strategy is gradually applied to mixtures of pseudocomponents, used to represent narrow crude oil fractions in process simulators. The limitations and the present need for the prediction of the composition of crude fractions on the molecular level will be discussed later. However, it is important to note here that the molecules in gasoline cuts can be fully analyzed by GS-MS. The analyses of gas oils is proceeding very fast with methods like integrated two dimensional GC, HPLC, supercritical GC, NMR, GC-MS, etc. [53, 54]. The analyses of heavier fractions with normal boiling points above 350°C and even up to 700°C+, is also rapidly developing [55]. A better evaluation of the experimental spectra, obtained by the analysis methods is achieved by modern statistical techniques often embedded in the respective devices [50-55].

Property values needed for product design

The property values for product design (performance property data) are even less than those for process design. In principle, the situation is better for fuel products, since many of the design properties – normal boiling temperatures, densities, viscosities, etc. determine also the performance of the compounds in formulated fuels. There is rather old work, done mainly in the US and the UK [56, 57], which is the basis of what we know about the octane numbers of around 200 gasoline compounds. It provides also blending numbers and values with the addition of tetraethyl lead. A database of the cetane numbers of diesel fuel relevant molecules has been recently published [58]. However, in both cases there is neither clear distinction between experimental and calculated values, nor reliability evaluation. The latter is especially important since the standard methods for estimation of these properties have changed significantly during the years. Properties of individual biodiesel molecules like lubricity, oxidation stability, cetane numbers, viscosity, etc., determined in standard tests can be found mainly in the works of Knothe and his co-workers [59-63].

Not surprisingly, there is an abundance of data for the performance of additives and components in gasolines and diesel fuels at different concentrations. While such data might be useful for deriving mixing rules, possible interactions with the unidentified compounds in the base fuels, diminish their value. The same holds for the development of predictive relationships between chemical structure and performance properties.

Most of the available data for properties of pure compounds, used in the specification of lubricants, are for viscosity. Often, but not always they can be used for calculations of viscosity indices. Data for lubricity, antioxidation effect and other important properties that can be quantitatively related to the chemical structure of pure additive compounds are very rare and rather old. Compatible experimental data from tribotests done in a well defined base oil (liquid paraffin), for instance, can be found in several publications only [64-68]. As with fuels most of the available data are for the performance of individual additives and their combinations in different base oils. The mechanism of action of the main classes of lubricant additives, however, is often well known, which allows for the development of semi-empirical relationships [25, 40]. There are also published (and more unpublished) data for the performance of different lubricant formulations in related tests [19, 23].

RELATIONSHIPS FOR PREDICTING PROPERTIES AND METHODS FOR DEVELOPING THEM

Computer aided design of fuels and lubricants requires the development of three main types of relationships: for prediction of properties of molecules from their chemical structure (“structure – properties”), for prediction of properties of mixtures of molecules with known structure and individual properties (“composition – properties”), and for prediction of the performance of fuel and lubricant formulations at a higher level of product screening from performance data estimated at the respective lower level (“properties – property”) [69-71].
Relationships for prediction of properties of molecules from their chemical structure

These relationships are applicable and widely used today for pure chemical compounds. They relate the numerical presentation of chemical structures to different physicochemical and performance properties. The first relationship of this type appeared after the Second World War [72]. The modern development, however, is due to their proven success in drug design [48]. By analogy with the name used in the design of biologically active molecules, these relationships are now known as “Quantitative Structure – Property Relationships, QSRRs” [73, 74] and used in the design of different chemical products [75].

For a given property, a QSRR can be schematically represented by the following equation:

\[ y_p = f(x_{p1}, x_{p2}, \ldots, x_{pk}, x_{pl}, x_{p2}, \ldots, x_{pm}; \beta_0, \beta_1, \ldots, \beta_n) \] (1)

where \( y_p \) is the property (e.g., boiling temperature, melting temperature, toxicity) to be predicted; \( x_{pi}, x_{p2}, \ldots, x_{pk} \) are descriptors representing numerically the molecular structure of the compounds in the database; \( x_{pl}, x_{p2}, \ldots, x_{pm} \) are known property data of the compounds in the database; and \( \beta_0, \beta_1, \ldots, \beta_n \) are the QSRR regression parameters [76].

It is seen from Eq. (1) that the development of a QSRR requires a database of molecular descriptors, presenting numerically the molecular structures of interest, and a database with experimental data for the same molecules. Availability of property data has been discussed in the previous section. Several approaches are available for numerical description of molecular structures and relating the latter to properties.

Group/bond contribution methods are the oldest and still the most common tool employed for the estimation of properties from molecular structure. They typically begin with decomposition of the molecular structure into particular groups and the counting of atoms in those groups (atom counts). Characteristic incremental values are assigned to the groups by regression of known data for the particular property. The whole molecule is then “restored” by summation of the contributions of all groups, and the property is determined as a function of the summations [46].

The group/bond contribution methods have become very popular because of their inherent advantages: explicit and comprehensive description of the molecules, straightforward restoration of the whole structure of the molecule, and simple calculation of the desired property. An important additional advantage is their direct application for generation of molecules having pre-designed properties [46, 77].

However, in the original group/bond contribution approach, the property of a compound is calculated from the number of occurrences of predefined fragments (groups) in the molecule. Since only the number of occurrences of a group and not the placement of the groups is considered, different isomers cannot be distinguished.

Higher order groups which account for some isomeric and cyclic structures and, thereby, increase the prediction accuracy have been proposed by Marrero and Gani [78] and recommended in [46]:

\[ f(X) = \sum_i N_iC_i + w\sum_j M_jD_j + z\sum_k O_kE_k \] (2)

where \( C_i \) is the contribution of the first-order group of type-\( i \) that occurs \( N_i \) times, \( D_j \) the contribution of the second-order group of type-\( j \) that occurs \( M_j \) times and \( E_k \) - the contribution of the third-order group of type-\( k \) that has \( O_k \) occurrences in a compound. In the first level of estimation, the constants \( w \) and \( z \) are assigned zero values because only first-order groups are employed. In the second level, the constants \( w \) and \( z \) are assigned unity and zero values, respectively because only first- and second-order groups are involved while in the third level, both \( w \) and \( z \) are set to unity values. The left-hand side of Eq. (2) is a simple function \( f(X) \) of the target property \( X \), which among other things “has to exhibit the best possible fit of the experimental data”.

The increased number of groups (and determined QSRR parameters) means statistically that more experimental data are needed in order to avoid chance correlations [79]. This is a serious problem [49], since, as already discussed, the number of available experimental data for most properties is low. The sophistication of the groups themselves has led to representations of the molecules that are rather complex [80].

The second most used today method for developing QSRRs has been described as the “significant common features, SCF” method [80]. Typically, this method aims at the selection, out of a huge descriptor database,
a small combination of descriptors, which comprise the most significant common features of the molecular structures, relevant to the particular property being targeted and describes adequately the available experimental data [33, 81 - 83]. The databases used today contain several thousand descriptors - groups/bonds description, topological indices, simulated molecular mechanics parameters, parameters from quantum mechanics simulation, etc. while the selected SCF combinations contain less than 10 descriptors.

The creation of descriptor databases is done with specialized software. There are many programs that can be used [33], among which probably the most popular in Europe is “Dragon” [84, 85] and in the US – “CODessa PRO” [86]. Both programs as an input use models of each targeted molecule, drawn with suitable software, for instance, “HyperChem” [87]. The files with the corrected 2D models can be used directly or 3D optimized models can be generated with the same program by simulated molecular mechanics, semi-empirical quantum mechanics methods, etc. Fig. 2 from [33] illustrates the main steps in the development of QSPRs, together with the methods for relating descriptors to properties.

For the development of the QSPR equations most of the available chemometric methods have been employed. Katritzky et al. [33] present a comprehensive summary of these. Among them multiple linear regression (MLR), multivariate non-linear methods, principal component regression (PCR), and partial least-squares regression (PLS), are the most often used, while artificial neural networks (ANN) and genetic algorithms are also gaining popularity. Some of these methods have been used for development of QSPRs relevant to fuel and lubricant design: MLR, PLS [33, 71, 80, 88], stepwise regression with orthogonal variables (SROV), principal component analysis (PCA) [89, 90], ANN [91]. They are employed also in the processing of the results for determination of composition, octane, cetane numbers and other fuel parameters by GC, GC-MS, NIR, NMR and other modern analysis techniques. Hybrid neural networks and genetic algorithms have been used also, in an exemplary published work of researchers from Purdue University and the “Lubrizol” company, for the development of injector cleaning fuel additives [91].

It is not a task of this work to discuss the application of modern statistical methods in detail, but it is important to note that methods with principal components need less data, and are suitable when the traditional regression methods cannot be applied, e.g. in QSPRs for additives. Hereunder, we shall concentrate on sharing our experience in the development of QSPR methods.

A QSPR is a relationship, which from the point of modern experimental design methodology may be classified as a passive experimental design [33]. Therefore, its predictive ability is strongly dependent on the variation of the features of the chemical structures contained in the property database. It has been demonstrated that variation of chemical structures and the principles of active experimental design can be applied to molecular structures [33, 71, 92, 93] for relatively small databases with structurally similar molecules. However, the practice nowadays is to employ all data available in commercial databases (i.e., DIPPR, TRC, etc.). As discussed previously, in these databases the chemical struc-

![Fig. 2. The main steps in the development of QSPRs.](image)
ture data are often contradictory and created at random. This leads to a typical problem of passive experiments – the “density” of the variation of structures (and that of the values of the independent variables) is not uniform, so the predictions are not uniform. For instance, if some structures are better presented in the database (there are more molecules of their type), the QSRR prediction for them is better than for structures which are worse represented. Moreover, the predictions for the compounds, used for derivation of the model (the so called “training set”) are better than for compounds, which are not used. This effect can be diminished by statistical means (e. g., creation from the database of alternative training and validation sets, and cross-validation) [71]. The randomness and the limited amount of available data also disprove often-made claims of some authors, that the QSRR they have developed can be applied outside the limitations of their property database. In fact, QSRRs are typical empirical correlations and it has been recognized in many of the published sources that any QSRR has an applicability domain, which is determined by the data used [71, 94].

The above problems have lead to the development of statistical measures of similarities between molecules (structures) [94]. One of the suggested approaches is in the heart of the “Quantitative Structure-Structure Property relationships, Q2PRs” [95, 96]. In Q2PR it is assumed that the vector of properties of the target compound \( t \) (the dependent variable) should be related to a set of \( m \) vectors of properties of predictive compounds (independent variables) \( p_1, p_2, \ldots, p_m \). The following partition of the \( t \) and \( p \) vectors to subvectors is used:

\[
t = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \quad p_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix}
\]

(3)

where \( x \) is an \( N \) vector of known properties, and \( y \) is a \( K \) vector of unknown properties. Both the \( N \) vector \( x \) and the \( K \) vector \( y \) contains known properties. In the Q2PR method, the subvectors \( x_1 \) and \( x_2 \) are molecular descriptors, while the subvectors \( y_1 \) and \( y_2 \) contain measured properties with various levels of experimental error. The structure-structure relationship, between \( x \) and the independent variables, \( x_1, x_2, \ldots, x_m \) is expressed by a linear regression model of the general form:

\[
x_i = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m + \epsilon
\]

(4)

where the weighing factors \( \beta_1, \beta_2, \ldots, \beta_m \) are the model parameters to be estimated and \( \epsilon \) represents independent normal errors of a constant variance.

The practical application of Eq. 4 allows for the identification of a small bank of up to 50 potential predictive compounds out of the much larger database. The same set of molecular descriptors must be defined for all compounds included in the database, while the span of the molecular descriptors should reflect the difference between the compounds in the database. Having the corresponding molecular descriptors for a target compound, \( x_i \) defined as well, a stepwise regression procedure is applied to the database in order to identify the most appropriate predictive compounds that should be included in the structure-structure regression model (Eq. 4), and to obtain the respective model parameters. Upon identifying the parameters of the structure–structure relationship model, the following equation can be used for predicting unknown properties, \( y_i \) of the target compound:

\[
y_i = \beta_1 y_1 + \beta_2 y_2 + \cdots + \beta_m y_m
\]

(5)

The properties that can be predicted for the target compound are all the properties for which experimental data are available for all the predictive compounds in the structure-structure correlation.

Q2PRs for compounds in homologous series are very valuable both from theoretical and practical point of view. Their properties can be determined with considerably higher precision by asymptotic behavior correlations (ABCs) and it has been suggested [33, 80] to use their predicted values as “pseudoeperimental” data in more diversified structurally databases and thus to improve QSPR models. The Q2PR approach has been used also for development of methods for homologous series [97, 98].

The Q2PR concept has been further applied in the development of the Targeted QSPR method [90, 99, 100]. The “targeted” QSPRs (TQSPRs) are more flexible in terms of similarity requirements. The method is based on preliminary identification of predictive compounds belonging to a similarity group of a particular target compound like in Q2PR. This similarity group
is then used to tailor a QSPPR for a particular property of the similarity group in the conventional manner (i.e., Eq. 1). The TQSPR method has also been proven to have advantages over conventional ABCs in reliable extrapolation of properties of homologous series [88, 101]. In this case, the TQSPR method identifies a single descriptor, which is collinear with the property values within the homologous series, so it requires less experimental data and can extrapolate linearly to high molecular mass members of the series.

The predictive ability of QSPPRs depends also on the structural information embedded in the used descriptor database. Usually, the latter contains a significant amount of correlated descriptors, so most of the software programs, used for their calculation, provide an option for excluding one descriptor of each correlated pair above a certain value of the partial correlation coefficient. However, most of the programs calculate topological descriptors, which use integer values for bonds and/or do not differentiate between “heavy” (i.e., except hydrogen) atoms – for instance, between 1-alkanols and n-alkanes with the same number of heavy atoms [101]. It has been suggested and demonstrated [33, 80], that the software for calculation of descriptors can use the real values (in angstroms) between the atoms of the minimized molecular models, which together with the identification of the heavy atoms in the diagonal of the matrix from which the topological descriptor is calculated, can avoid the identified problems. The reviewed above QS2PR methods have demonstrated superior prediction quality, as compared to conventional QSPPR techniques. The main explanation for their success is that they are developed for structurally similar compounds and better reflect their significant common features. A huge structurally diversified database is more similar to a “black box” in which significant portion of valuable information is lost.

Relationships for prediction of the properties of mixtures

The review of Katritzky et al. [73] gives examples of the use of QSPPRs for predicting properties relevant to process design, as well as performance properties that can be useful in fuel and lubricant design of individual compounds – flash temperatures, octane numbers, autoignition temperatures and cetane numbers, etc. Thus, they can be and are successfully applied for identification of molecular structures with specified properties for individual compounds, used as additives [91], synthetic components of fuels and lubricants, etc.

For mixtures of discrete components, i.e. with known chemical structures and known number of components, the QSPPRs can be combined with semi-empirical mixing rules, containing interaction coefficients, determined from experimental data [46]. This approach is well developed for process design and implemented in process simulators. In the latter, the mixing rules and the interaction coefficients are selected as components of the thermodynamic package. It has been demonstrated that the TQSPR method can be used for estimation of binary interaction coefficients of homologous hydrocarbons for the Soave-Redlich-Kwong and Peng-Robinson equations of state [102]. In the same paper new methods, based on QS2PR, are suggested also for calculation of temperature dependent parameters.

For continuous mixtures with unknown number of components like the crude oil fractions, the system has to be divided into physical pseudo-components and their physicochemical properties have to be defined. Modeling of processes with chemical transformation requires chemically defined lumps (e.g., hydrocarbon groups, etc.). In the ideal case, all or most of the data would have been experimental. Unfortunately, this requires detailed time-consuming analyses, which is not always affordable in a refinery (e.g. when monitoring the technological regime of an installation). The minimum of experimental data, required by process simulators is usually for the TBP and density distribution curves of the feed. The rest of the data can be estimated by different methods. Probably, the most widely used methods are those for estimation of the distribution of properties in crude oils, described comprehensively and with many examples by M. Riazi [28, 103]. Riazi provides also data for properties of pure hydrocarbons, compiled by the American petroleum institute (API) up to 1977, and correlations, developed from them. We are not going to discuss these correlations here, since the API data are rather old and some of them are not experimental.

Riazi’s generalized correlation for estimation of boiling point distribution and producing TBP curves from assay data is well known. This correlation, as well as correlations for distribution of other properties, have been validated and used by Bulgarian scientists [104]. As an illustration, hereunder, we present one less known generalized correlation for distribution of properties [28]:

227
\[ \theta = \theta_m - \exp(a - bM^c) \]  
(6)

where \( \theta \) is the calculated property (NBP, density, refractive index, critical temperature, pressure, density, acentric factor, solubility parameter or surface tension), \( M \) is the molecular mass and \( \theta_m \) is the limiting value of the property when \( M \to \infty \). Tabulated values for the constants \( a, b, c \) and \( \theta_m \) are tabulated in [28]. The value of \( M \) is calculated from the mean average boiling temperature of the fraction and is the only experimental data used.

Correlations for estimation of the distribution of chemical groups, sulfur, nitrogen, etc. in crude oil fractions from their physicochemical properties, proposed by Riazi and other authors are also available [28, 103].

Distribution correlations are in fact empirical or semi-empirical “property – property” correlations, since they are developed from known values for the properties of fractions from different crude oils to calculate unknown properties of fractions of from the same or another crude oil. Process simulators are supplied with lists of available methods of the above type, among which the user can select. The values they calculate are then used by the thermodynamic package for the particular simulation.

Property – property correlations of the same type are available also for many performance properties of continuous hydrocarbon mixtures used as components and/or base fuels and oils – flash and pour point, octane and cetane numbers, cetane index, viscosity, aniline point, etc. [28].

When the fractions are used as components, blending rules are also required. Some of the properties of the blends are additive (e.g., density, molecular mass, sulfur and nitrogen content, etc.). However, many important properties (octane and cetane numbers, Reid vapor pressure, viscosity, smoke point, etc.) are not additive properties. Octane numbers are additive for \( n \)- and \( iso \)-alkanes, but interactions (synergism or antagonism) between other groups of hydrocarbons might lead to \( \pm 8 \) units of octane number deviations [105]. In order to go around that problem modern refineries use the so called “blend indices” or “blending numbers”. These are calculated by empirical property – property correlations, similar to the described above. For instance, the flash point blend index of a component, \( BI_{FP} \) [103], can be calculated by:

\[ BI_{FP} = FP_{FP} \]  
(7)

where \( FP_{FP} \) is the flash point of the component \( i \) and the best recommended value for the constant \( x \) is -0.06. The flash point of the blend is then calculated additively:

\[ BI_{FP, Blend} = \sum_{i=1}^{n} x_i BI_{FPi} \]  
(8)

where \( x_i \) is the volume fraction of component \( I \), \( n \) is the number of components, the blend index of each component is calculated from Eq. (7).

The deviations of linearly calculated values from those calculated with blend indices can be significant. Al Sahlaf et al. in [28] give an example for a research octane number deviation of 3 octane units, which is not acceptable.

The correlations and mixing rules for crude oil fractions discussed in this section, are comprehensive and easy to work out; they use data from assays and routine analyses; they are available in process simulators, etc. However, their constants are tied to the fractions of crude oils used in their derivation and for different crude oils may have huge deviations. Moreover, the subsequent use of calculated values transfers the errors from one correlation to the next and eventually – they are accumulated in the final results. Typically, correlations for crude oil fractions should not be used for individual compounds and might not be applicable for fuels and oils containing alternative components, derived from coal, natural gas, renewable raw materials, some additives, etc.

Their inferiority as compared to molecular correlations has been demonstrated. For instance, in [106] it has been shown that for simulation of a vacuum distillation column, separating a model mixture of three individual hydrocarbons, it might not be possible to determine the tray for drawing the middle component correctly, when applying different distribution property methods for estimation of critical temperatures of the pseudo-components, provided in a process simulator. Our unpublished results [107] indicate that in the simulation of atmospheric distillation with a fore-column and a main column, separating 4.8 mln t/year of crude oil, represented by a model mixture of hydrocarbons, the errors in the prediction of the yields of the fractions, due only to the use of distribution property methods, provided in the simulator, can be significant. Al-
ternatively, in both cases [106, 107], a method based on structure-property estimation of critical temperature, like those discussed in the previous section, determines correctly the number of the tray and with an acceptable error - the yields of the fractions.

One of the correlations for calculation of octane numbers of petroleum fractions, which uses non-linear mixture rules, similar to those used in thermodynamic packages, has been proposed by Twu and Coon [105]. The authors have estimated the interaction parameters with experimental data for the concentration of alkenes, arenes and alkanes in 161 gasoline cuts, obtained with different processes from different crudes, and their motor (MON) and research (RON) octane numbers. The Twu and Coon correlations is recommended as one of the methods for calculation of octane blend indices in refineries [28]. For the original data, used in the development of the model, the MON correlation is shown to predict the experimental data with an average error of 1.19 % [105]. Figure 3 shows our results for the prediction of experimental MON data of 223 binary and ternary mixtures of individual hydrocarbons (taken from [56]) by the original Twu and Coon correlation [105].

The results from Fig. 3 confirm the theoretical assumption that even the more sophisticated correlations, developed with data only from crude oil cuts, cannot be used for individual hydrocarbons, and thus

![Graph showing prediction error](image)

**Fig. 3.** Errors in prediction of motor octane number (MON) by the Twu-Coon correlation.

...luck consistency. In summary, if better prediction is needed, the property distribution methods have to be replaced by correlations with mixing rules, developed from experimental data for mixtures of molecules. This is already possible for mixtures of methyl esters of fatty acids (biodiesel), the molecular types in which are less than 10, and their concentrations can be analyzed by HPLC. Properties for individual molecules without measured values can be estimated with QSPRs, like described in the previous section. The same approach is feasible for additive packages and synthetic components.

### Table 3. Properties that can be predicted with the new methods and their use in simulation.

<table>
<thead>
<tr>
<th>Process design</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimation</strong>: NBP, Critical properties, Density, Vapor pressure, Viscosity, Surface tension, Acentric factor, Activation energy, Activity coefficient, Enthalpies of formation and reactions, Reaction constants, Solubility parameters, Polymer free volume, etc.</td>
</tr>
<tr>
<td><strong>Simulation</strong>: Phase equilibria (Distillation, Adsorption, Extraction, Absorption, Reactive distillation, Crystallization, Membrane separation, Mixing, Free evaporation, etc); Adsorption and catalytic reactions; Reaction networks and mechanisms (in Pyrolysis, Cracking, Reforming, etc.); Design of benign solvents; Environment pollution (Spreading of pollutants in air, water, soil, etc.)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Product design</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimation</strong>: Properties (ON, CN, RVP, Viscosity, Distillation curves; Melting, pour point and flash temperatures, Ignition delay, Friction coefficients, seizure and weld loads in tribotests, etc); Results from performance tests (Injector deposits, induction period and other parameters of oxidation, etc.)</td>
</tr>
<tr>
<td><strong>Simulation</strong>: Action mechanism (e.g., In Lubrication: adsorption, chemisorption, surface reactions; Reaction networks (e.g., In Oxidation, Ignition and Combustion); Chromatographic and Spectral properties (e.g. spectra and retention times, etc. in GC, MS, HPLC, IR, UV, NMR, etc.); Synergism between additives, components, etc.; Molecular design of additives and components, etc.)</td>
</tr>
</tbody>
</table>

229
For crude oil cuts a more evolutionary approach seems more appropriate, as will be discussed later.

The new methods can be used for very diversified tasks. Two examples are the complex phase equilibria modeling, relevant both to upstream and downstream crude oil processing [108] and the determination of constants in semi-empirical modeling of lubrication [71]. Table 3, abbreviated and adapted from the original table in [109] summarizes properties, which the new molecular simulation methods, described above can predict and how they can be used in process and product design with discrete mixtures.

**Synergism and antagonism and “properties – property” relationships for fuels and lubricants design**

*Achieving synergism and avoiding antagonism* is a major aim in the development of any additive product, because synergism saves money. As already discussed, it can be achieved by non-linear mixing rules and by correlations developed by active design experiments for mixtures. In a mixture with fixed total amount of components, synergism is observed if there exists a combination with functional property value higher than that of the best single component, tested in concentration equal to the total amount of components used. Fig. 4 illustrates the effects observed in the formulation of a mixture of two fuel or lubricant additives.

As discussed above, the effects shown on Figure 4 can be identified experimentally or if an adequate relationship is already available and it is used for optimization of the composition, with the goal to find the maximum of the TE, if it exists. For more than two components, the typical response surface method used to solve this more general chemical engineering problem is the active design experiment. Bulgarian scientists have contributed both to the theory of responsive surface methodology and to its use in fuel and lubricant design. These methods can be also supplemented by QSRRs for determination of molecules, which can potentially provide greatest contributions to the studied mixtures.

Since, fuels and lubricants are developed to achieve pre-set values of different properties, defined in a specification, and they all depend on concentration, the formulator has to solve also a reconciliation problem, because it is rare to achieve synergism for all properties [110]. It is possible also to solve numerically the reconciliation problem. The solution is in the form of ranges of concentrations of the components in which the mixture displays synergism, positive effects or at least indifference for all studied effects.

*“Properties – property” relationships for fuels and lubricants design* are needed in order to minimize “try and fail” experiments in the development of fuels and lubricants [70]. These products are evaluated firstly by express methods which are relatively cheap, do not require huge samples, and provide quickly approximate, but comprehensive assessment of the expected performance of the alternative formulations in real service. Because of these features, they can estimate a significant number of alternative formulations at the laboratory level. They are also used for control in the production of fuels and lubricants. Examples of such test for fuels are the simulated distillation, estimation of octane numbers from GC and NIR data, estimation of cetane numbers from autoignition delay, etc. For lubricants such tests estimate shear resistance of viscosity additives, lubricating properties on simple tribometers, antioxidation stability, corrosion resistance, etc.

As shown in this review, there are already relationships, which can relate the molecular structure (e.g. of additives and components) to their performance in
such tests. Moreover, there are relationships for process design, which can provide pre-set composition and properties of the production streams, and their blends. Methods for petroleum fractions and blended base stocks can be developed further, but even now they perform adequately within the present needs for precision.

In the development algorithm the express tests select a smaller number of formulations, which can eventually satisfy the requirements of more sophisticated and expensive pilot tests. The tests for lubricants and especially – for automotive oils, are most complex and their development passes through a number of levels of increased sophistication and relevance to real service.

Modern “properties–property” methods have to establish relationships between the different testing levels in the form of correlations in which lower level properties are independent variables and the dependent variable is a higher level property [70]. Figure 5 shows how these relationships can be combined with the relationships discussed above in an algorithm for systematic design of lubricants.

CONCLUSIONS

In the introduction, the major challenges facing the petroleum processing industry have been outlined. Then new methods which are being developed in the leading laboratories of the private and public sector, and their relevance to the new challenges were presented. The presently used methods of the petroleum industry and their problems were analyzed vis-a-vis the opportunities, which the new methods for process and product design of fuels and lubricants can open up. However, the review also identified challenges that face the implementation of the new methods, which are discussed in essence below.

One important problem is the need for more validated property data for the molecules of interest to the petroleum industry, including additives and components. Collection of more data for fractions from different crudes, BTL, CTL, F-T gasols, etc., is also needed. One possible solution might be the organization at national and/or international level of efforts similar to the API Projects [56, 57].

The second problem concerns collection of data from evaluation of fuel and lubricant formulations at different levels of the development algorithm (Fig. 5). Since, the methods with which they are obtained are standardized, such data are already available in company databases, and more will be accumulated. Moreover, these data are implicitly related because the simpler tests (e.g., express and full time oxidation, simple tribometers and tests like the FZG machine, L-37 and L-42, etc.; express corrosion tests by electrochemical methods and corrosion tests in chambers, etc.) have
all been designed to model real life applications. One objection against publishing such data might be the need to protect company secrets. However, there is no need to reveal the composition of the particular fuel or lubricant. What the development of relations between tests requires is the results from lower level and higher level tests; the samples may be identified only with numbers [70].

The precision of models cannot be higher than the experimental precision of the data, used in their derivation. The chemical engineer needs to know precision of the calculated data and the limitations of the correlations, in order to compare them with the requirements of potential applications. From this point of view, the presently used distribution of properties methods for fractions will continue to be used, as long as they satisfy the precision expected from them. Molecular level methods seem impractical for fractions at least for now, even if all molecules can be identified within reasonable time and effort. However, predictions can be greatly improved, if new non-linear correlations are derived with mixtures of individual molecules and fractions [111].

The challenges for the petroleum industry, outlined in the beginning, are piling up and the need for implementation of information-based technologies is pressing. Methods for presentation of fractions and even crude oils by composition matrices and methods for characterizing them down to molecular level are improving. New devices for in-line analyses are offered [32, 36, 53, 55]. Therefore, the new methods are coming. Derived from and validated against more data with innovative statistical techniques, they will make upstream and downstream processing more competitive, and more ecologically compatible.

REFERENCES

45. An internationally recognized nomenclature system for petroleum additives, ATC document 31, Technical Committee of Petroleum Additive Manufac-
70. G.S. Cholakov, Some approaches to computer aided design of lubricants, Oxidation communications, 17, 1994, 4, 303-310.
71. G. Cholakov, K. Stanulov, P. Devenski, H. Iontchev:


