

Article

Visual Analysis of Odor Interaction Based on Support Vector Regression Method

Luchun Yan ¹, Chuandong Wu ² and Jiemin Liu ^{2,*}

¹ School of Materials Science and Engineering, University of Science and Technology Beijing, Beijing 100083, China; yanluchun@126.com

² School of Chemistry and Biological Engineering, University of Science and Technology Beijing, Beijing 100083, China; dong1035@126.com

* Correspondence: liujm@ustb.edu.cn

Received: 18 February 2020; Accepted: 16 March 2020; Published: 19 March 2020



Abstract: The complex odor interaction between odorants makes it difficult to predict the odor intensity of their mixtures. The analysis method is currently one of the factors limiting our understanding of the odor interaction laws. We used a support vector regression algorithm to establish odor intensity prediction models for binary esters, aldehydes, and aromatic hydrocarbon mixtures, respectively. The prediction accuracy to both training samples and test samples demonstrated the high prediction capacity of the support vector regression model. Then the optimized model was used to generate extra odor data by predicting the odor intensity of more simulated samples with various mixing ratios and concentration levels. Based on these olfactory measured and model predicted data, the odor interaction was analyzed in the form of contour maps. This intuitive method showed more details about the odor interaction pattern in the binary mixture. We found that the antagonism effect was commonly observed in these binary mixtures and the interaction degree was more intense when the components' mixing ratio was close. Meanwhile, the odor intensity level of the odor mixture barely influenced the interaction degree. The machine learning algorithms were considered promising tools in odor researches.

Keywords: odor intensity; odor evaluation; machine learning; prediction model; antagonism effect

1. Introduction

As one of the important environmental pollution problems, odor pollution will cause people potential health hazards and uncomfortable feelings. Many countries have published related laws and regulations about the emission limits in conventional pollution sources like chemical industries, livestock processing industries, and sewage treatment systems [1]. Although, sometimes the exhaust gas has met the required emission limits but its odor is still very obvious. This phenomenon mainly comes from the superposition effect and other possible interactions caused by the mixing of multiple low-concentration odor pollutants [2]. In these cases, the olfactory evaluations by specialized human assessors are often used [3]. However, the test cost (e.g., a panel of specialized assessors, professional testing laboratory, long sampling, and testing cycles) of olfactory evaluation is much higher than regular chemical analysis [4]. Therefore, the relationship between the chemical composition and the odor intensity of odor mixtures have been widely researched in related fields [5]. The odor intensity prediction models and electronic nose are also urgently needed for both research and application purposes [6].

For individual odorants, the Weber–Fechner law and Power Law model have been widely used for the conversion of chemical concentration and odor intensity [7]. The linear relationship between odor activity value (ratio of chemical concentration to its odor threshold) and odor intensity is also proposed

for odor intensity prediction [8]. For odor mixtures, many models like the ERM model, Vector model, U model, Additivity model, and their extended versions are reported [9–11]. These models provide valuable ideas and guidance for our understanding of the odor interaction. Sometimes, they also can be used in odor intensity prediction. In general, their prediction accuracy and applicable scope are often limited [12]. Even though some models have achieved accurate odor intensity prediction, the key parameters in the model must need to be experimentally measured for target substances [13]. It severely limits the convenience and practical application of these methods. Besides, these empirical models mainly focus on the mathematical associations among the target variables. The display of odor interaction law is abstract, and its feasibility to simulate and analyze influencing factors is weak [14]. Therefore, researchers are always trying to find more effective methods to investigate odor interaction. For instance, Teixeira et al. applied the perfumery ternary diagram (PTD) and perfumery quaternary diagram (PQD) methodologies to map the predicted odor intensities of fragrance mixtures [15].

In recent years, machine learning methods have developed rapidly in their algorithms and implementation techniques. Beyond traditional computer science, machine learning methods are increasingly being used for scientific researches in fields like chemistry, materials, and biology [16]. There are two important advantages of machine learning: excellent data analysis and processing capabilities to both the big data (a large amount of data and variables) and the small sample dataset (which is also a typical feature of odor data) and functional generalization capacity to new samples [17]. For example, Arabgol et al. used the support vector machine and 160 water samples data to build a nitrate concentration prediction model, and it successfully predicted the map of nitrate concentration for all four seasons [18]. Besides, machine learning also provides rich visual data analysis and utilization techniques [19]. These advantages are precisely the functions that we want in the odor interaction research. Machine learning methods have also attracted the attention of odor researchers and have been applied in various forms [20]. Szulczynski et al. proposed an electronic nose and the odor intensity was directly linked to the results of analytical air monitoring with a fuzzy logic algorithm [21]. Zhu et al. made an accurate prediction of soil organic matter contents by employing back-propagation neural network, support vector regression, and partial least squares regression methods [22]. Thus, we think that machine learning methods can also provide more useful tools in the study of odor interaction.

In this paper, odor data of binary esters, aldehydes, and aromatic hydrocarbons mixtures were collected from our previous studies. The support vector regression algorithm was employed to establish the odor intensity prediction model which achieved the direct conversion from the mixture's chemical composition to its odor intensity. The optimized model was used to produce more olfactory evaluation data of similar odor mixtures. Based on these odor data, we proposed a visual analysis method of odor interaction. The influences of the components' mixing ratio and the sample concentration level on odor interaction were investigated intuitively. With the help of machine learning methods, we hope to find more effective and intuitive analysis methods of odor interaction.

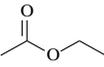
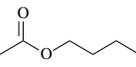
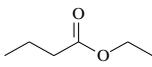
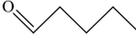
2. Materials and Methods

2.1. Stimuli and Odor Data

As typical odor pollutants in odor sources like landfills and sewage treatment plants, odor data (i.e., odor threshold, Table 1; olfactory measured odor intensity and chemical concentration) of ethyl acetate (EA), butyl acetate (BA), ethyl butyrate (EB), propionaldehyde (PA), *n*-valeraldehyde (VA), *n*-heptaldehyde (HEP), benzene (B), toluene (T), ethylbenzene (E), and some of their binary mixtures were collected from our previous studies [23–25]. In these experiments, the odor samples were prepared through transferring a certain amount of standard gas to an odor-free plastic bag (3 L volume; Sinodour, Tianjin, China) and diluted with purified air. A sensory panel (8–14 human assessors) was employed to measure the odor threshold of each stimulus and the odor intensity of odor samples. The ASTM odor intensity referencing scale (OIRS, water solutions of 1-butanol from level 1 (aqueous solution of 12 ppm) to level 8 (1550 ppm) with a geometric progression of two at 27 ± 1 °C) was used as the

standard in odor intensity evaluation [26]. More details about the testing procedure and environmental requirements were described in these references. In this study, the collected dataset contains 31 samples of binary mixture EA+BA, 21 samples of EA+EB, 22 samples of BA+EB, 24 samples of PA+VA, 24 samples of PA+HEP, 24 samples of VA+HEP, 34 samples of B+T, 31 samples of B+E, and 24 samples of T+E.

Table 1. List of odorants and their odor thresholds.

| Order | Odorant (Abbreviation) | CAS# | Chemical Structure | Odor Threshold/mg/m ³ |
|-------|------------------------------|----------|--|----------------------------------|
| 1 | ethyl acetate (EA) | 141-78-6 |  | 0.276 ^I |
| 2 | butyl acetate (BA) | 123-86-4 |  | 0.085 ^I |
| 3 | ethyl butyrate (EB) | 105-54-4 |  | 0.053 ^I |
| 4 | propionaldehyde (PA) | 123-38-6 |  | 40.6 E-3 ^{II} |
| 5 | <i>n</i> -valeraldehyde (VA) | 110-62-3 |  | 20.5 E-3 ^{II} |
| 6 | <i>n</i> -heptaldehyde (HEP) | 117-71-7 |  | 26.0 E-3 ^{II} |
| 7 | benzene (B) | 71-43-2 |  | 2.53 ^{III} |
| 8 | toluene (T) | 108-88-3 |  | 1.43 ^{III} |
| 9 | Ethylbenzene (E) | 100-41-4 |  | 0.45 ^{III} |

^I Odor detection thresholds in reference [24]; ^{II} Odor detection thresholds in reference [26]; ^{III} Odor detection thresholds in reference [25].

2.2. Support Vector Regression Methodology

Rooted in statistical learning or Vapnik-Chervonenkis (VC) theory, support vector machines (SVMs) are well positioned to generalize on yet-to-be-seen data [27]. The SVMs are very popular and effective in solving classification problems. The SVMs algorithm aims at establishing a clear gap (in the form of a line/hyperplane) that is as wide as possible to divide the sample points in a certain space. For non-linear separable data, a kernel trick is employed to map these data into high-dimensional feature spaces where the data can be successfully separated by a hyperplane. The parameters of the SVMs model will be optimized to find an optimal hyperplane that maximizes the margin of the decision boundary. The Support Vector Regression (SVR) uses the same principles as the SVMs, except that the hyperplane optimization focuses on covering as many data points as possible within a fixed-width boundary [28]. In the SVR model, the radial basis function (RBF) is mostly chosen as the kernel function, which is demonstrated quite effective in transforming non-linear data [29]. As two key parameters influencing the RBF kernel, *C* is to control the punishment degree of sample error, and γ is whether the accuracy is allowed to be greater than or equal to 1 for the samples of misclassification [30]. The working mechanism, mathematical formulas and optimization strategies of the SVR model have been abundantly reported in the literature [31]. Generally, the SVR model has high accuracy and excellent generalization capacity for small sample data. These advantages perfectly match the data characteristics and analysis requirements of common odor data.

In this study, the collected dataset of each binary mixture was randomly divided into two parts: the training set (70% amount) and the test set (30% amount). The training set was used in the model

training and optimization steps. The chemical concentration value of each substance was used as the input variables of the SVR model, and their mixture's olfactory measured odor intensity value was the model's target output variable. The grid search scheme was used to search the model's hyperparameters (the RBF kernel was used, hyperparameters C and γ were optimized), and the 10-fold cross-validation was used to evaluate the prediction ability of a model with certain hyperparameters [32,33]. Based on an optimized SVR model, the odor intensity value of a mixture can be directly predicted after inputting its corresponding chemical composition to the model. Finally, the model's predictive capacity to new data would be verified by the test set. All of the above statistical analysis and data mining work was conducted using Python software and scikit-learn toolkit. For the main modeling and simulation content in this study, we uploaded the corresponding Python code in ipynb file format as the Supplementary Materials.

2.3. Experimental Procedure

This study firstly established SVR models for the odor intensity prediction of binary mixtures of esters, aldehydes and aromatic hydrocarbons. After the SVR model parameters were optimized, its predictive performance was verified by comparing the olfactory measured odor intensity with the SVR predicted odor intensity values. To evaluate the accuracy and precision of the SVR model, the coefficient of determination (R^2) and the mean absolute error (MAE) were individually calculated for training samples and test samples. The optimized SVR models were then used to predict the odor intensity for other similar odor mixtures, and then more odor data was produced. Based on the accumulated odor data, the following strategy was used to explore the odor interaction in a more intuitive way. Here, the odor intensity value of individual substance (OI) was calculated on the basis of our previous obtained OI - $\ln OAV$ (natural logarithm of odor activity value) equations [24–26] and odor threshold values ($C_{thr.}$) in Table 1:

$$\ln OAV = \ln\left(\frac{C}{C_{thr.}}\right) \quad (1)$$

For esters:

$$OI = 1.40 \cdot \ln OAV - 2.70 \quad (2)$$

For aldehydes:

$$OI = 1.76 \cdot \ln OAV - 1.82 \quad (3)$$

For aromatic hydrocarbon:

$$OI = 1.07 \cdot \ln OAV \quad (4)$$

For binary mixtures without odor interaction, its ideal odor intensity was defined as $OI_{sum.}$:

$$OI_{sum.} = OI_a + OI_b \quad (5)$$

where, the value of OI_a/OI_b was calculated on the basis of its chemical concentration in the mixture (Equations (1)–(4)). Both the olfactory measured odor intensity and the SVR predicted odor intensity of binary odor mixture were all marked as $OI_{mix.}$, and we defined the odor interaction degree in the mixture as OI reduction and OI reduction ratio:

$$OI \text{ reduction} = OI_{sum.} - OI_{mix.} \quad (6)$$

$$OI \text{ reduction ratio} = \frac{OI_{sum.} - OI_{mix.}}{OI_{sum.}} \quad (7)$$

Both *OI reduction* and *OI reduction ratio* were considered as indicators of interaction degree in the binary mixture. Based on these two variables, their relationships with the component's concentration and components' mixing ratio were carefully investigated. The contour maps and scatters plots were employed to display the odor interaction pattern of binary mixtures intuitively.

3. Results and Discussion

3.1. Odor Intensity Predictive Performance of the SVR Model

As shown in Figure 1, the olfactory measured odor intensity and the SVR predicted odor intensity were compared in the form of scatter plots. When the sample point was close to the diagonal (red line), it meant that the SVR model made an accurate prediction. For most of the training samples and test samples, the SVR models successfully made perfect predictions. It demonstrated the feasibility and good fitting ability of the SVR algorithm in regular odor data analysis. Besides, the similar predictive accuracy between training samples and test samples proved that these SVR models were not overfitted. The overfitted model usually will correspond exactly to the training set, and may, therefore, fail to predict future observations reliably (like the test sample). This phenomenon is usually caused by the strong fitting ability of machine learning algorithms and its improper parameter settings, which is one of the key issues that should be avoided in the application of machine learning methods [34]. Table 2 listed the coefficient of determination (R^2) and mean absolute error (*MAE*) between olfactory measured odor intensity and SVR predicted odor intensity of each odor mixture individually. From the R^2 values of training samples and test samples, it also confirmed that the SVR models had good predictive accuracy and it was not over-fitted. Different from the other mixtures, the R^2 values of mixture T+E was lower. It probably was caused by a relatively poor accuracy of the olfactory measured results. Because the SVR algorithm is very sensitive to the noise in the training data [35]. Therefore, the noise (arising from the error of olfactory evaluation) in the training samples can easily affect the fitting effect of the SVR model. Nevertheless, the *MAE* results still showed that the prediction error of the SVR models was very limited. In the regular olfactory evaluation tests, the 0.4 OIRS level of error was usually observed and widely accepted [25]. Thus, the optimized SVR models were considered to be useful and accurate in the odor intensity prediction of these binary odor mixtures.

Table 2. The prediction accuracy of SVR models. The coefficient of determination (R^2) and mean absolute error (*MAE*) were individually calculated for each kind of binary mixture.

| Mixture | R^2 | | <i>MAE</i> | |
|---------|--------------|----------|--------------|----------|
| | Training Set | Test Set | Training Set | Test Set |
| EA+BA | 0.97 | 0.95 | 0.15 | 0.26 |
| BA+EB | 0.96 | 0.85 | 0.15 | 0.33 |
| EA+EB | 0.87 | 0.87 | 0.17 | 0.25 |
| PA+VA | 0.95 | 0.78 | 0.14 | 0.25 |
| PA+HEP | 0.96 | 0.94 | 0.17 | 0.23 |
| VA+HEP | 0.97 | 0.87 | 0.15 | 0.31 |
| B+T | 0.87 | 0.81 | 0.33 | 0.43 |
| T+E | 0.78 | 0.68 | 0.31 | 0.40 |
| B+E | 0.98 | 0.94 | 0.09 | 0.27 |

The odor intensity prediction models were considered promising techniques in the field of odor evaluation. First, the prediction models could directly perform the odor intensity evaluation (basing on the composition and concentration information measured by analytical equipment) instead of human assessors. The influences of assessor quantity, age, gender, and testing environments could be avoided [36]. On the other hand, it has been reported that the e-nose can directly perform odor intensity evaluation [21,37]. However, it directly correlates the sensor signal and odor intensity, and does not fully consider the gas mixture's composition. Therefore, the device is more focused to specific

target gases. In contrast, e-noses and online monitoring devices capable of gas identification and concentration detection are more common and more mature [38,39]. If combining the odor intensity prediction model with these e-noses and online monitoring devices, it will significantly improve their olfactory assessment capacity and extend the applicable scope.

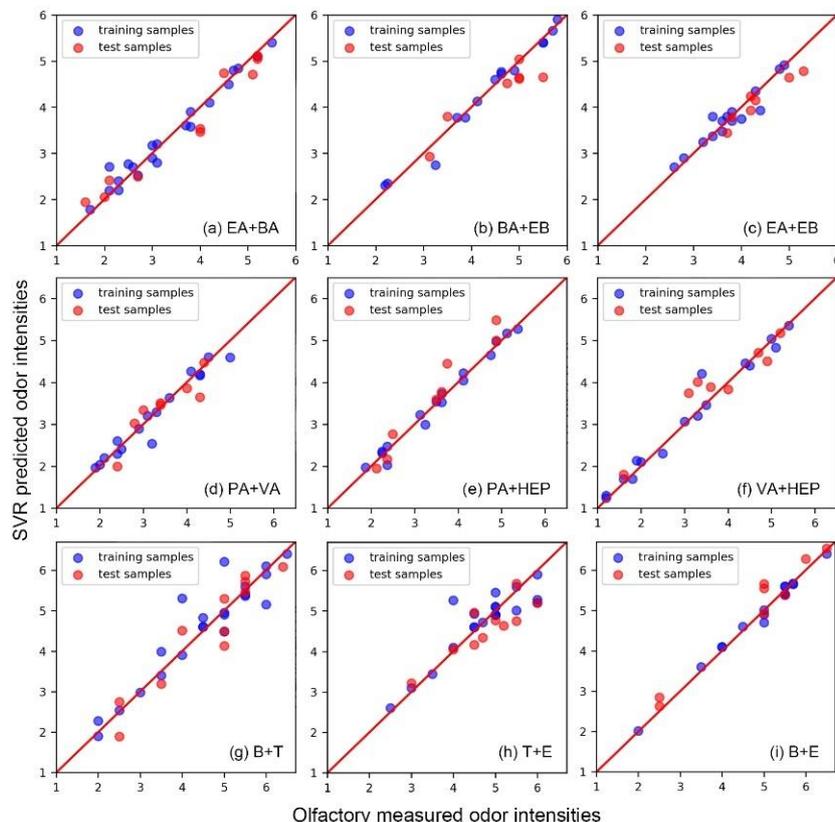


Figure 1. The comparison between olfactory measured odor intensity and Support Vector Regression (SVR) predicted the odor intensity of nine different binary mixtures.

3.2. SVR-Assisted Visual Analysis of Odor Interaction

In comparison with traditional odor intensity prediction models, some machine learning methods like SVR have a significant advantage and also one of its disadvantages. The mechanism like a black box severely limited its function in explaining related mechanisms and laws. Although many studies have established empirical models to explain the odor interaction phenomenon and made conclusions, we still hope to develop more analytical methods through the reasonable use of machine learning methods. Since it has been proved that there is a simple linear relationship between OI and $\ln OAV$ of an individual substance, we think that using the $\ln OAV$ value to represent the content of a component is also helpful in odor interaction studies [23,24]. As illustrated in Figure 2a, c, and d, scatter plots of the relationship between each component's content (in the form of $\ln OAV$) and the odor interaction degree (in the form of OI reduction values; i.e., the color of each dot) was plotted. Based on the definition of OI reduction in Equation (6), the larger OI reduction value meant the stronger degree of antagonism effect [40]. It could be seen that when the content of both components was high, the antagonism effect would be more intense. Because the amount of actual olfactory measured odor data was limited, this scatters plot only provided little information and the results were not intuitive enough.

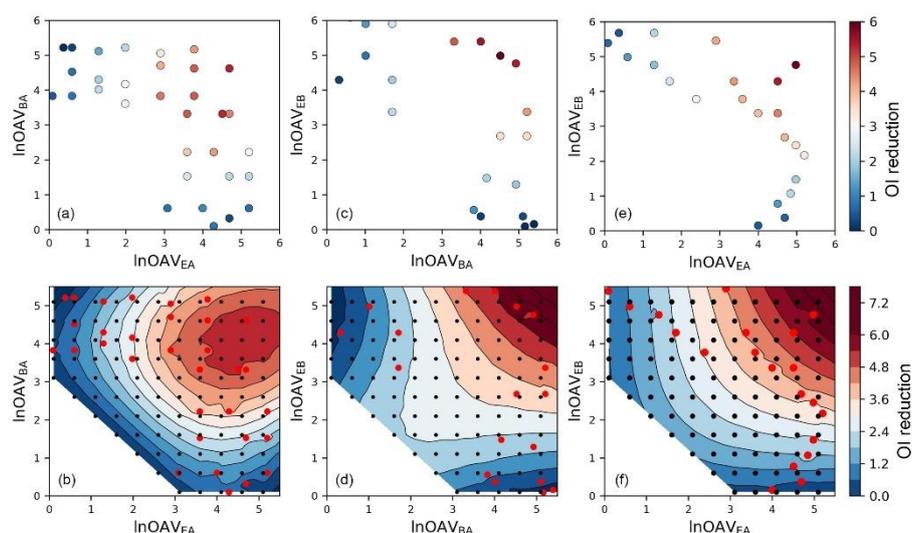


Figure 2. The scatter plot of the relationship between olfactory measured odor intensity reduction degree (*OI reduction*) and components' $\ln\text{OAV}$ (natural logarithm of odor activity value) values, and corresponding contour map after adding SVR model predicted data (in the contour map, red dots: olfactory measured data; black dots: SVR predicted data) for binary mixture (a,b) EA+BA, (c,d) BA+EB, and (e,f) EA+EB.

In order to obtain more odor evaluation results, we used the SVR model instead of olfactory measurement which saved much time and economic costs. Since the core idea of machine learning was to find out the mathematical relationship between the chemical composition and the corresponding mixture's odor intensity, a certain amount of training samples usually could guarantee the modeling effect. After that, the optimized model would also be valid for other similar samples. This strategy and function have been fully demonstrated and widely applied in many research [41]. As shown in the above R^2 and MAE analysis results of the training and test samples, the SVR model had obtained the correct mapping relationship. In this case, the optimized model also will be valid for other samples of the corresponding mixture with different chemical concentrations. As shown by the black dots plotted in Figure 2b,d,f, we predicted the odor intensity of many binary mixtures with different chemical contents. Results from actual olfactory measurements in our previous studies were plotted with red dots. In order to distinguish the data source here, the color of the dots no longer indicates the *OI reduction* value like Figure 2a,c,e. Based on these olfactory measured and SVR predicted results, the contour maps about the *OI reduction* degree and mixtures' composition were plotted. Through these diagrams, the interaction of odor substances became more intuitive. It could be concluded that the degree of antagonism effect was usually weak if the $\ln\text{OAV}$ value of any component is low. When the content of one component was constant, the antagonism degree would increase as the content of the other component enhancing. Besides, the antagonism degree would become more intense when the $\ln\text{OAV}$ values of the two components were approaching close.

In the same way, the odor interaction of binary aldehydes mixtures was also analyzed (Figure 3). The interaction pattern of binary aldehydes mixtures was almost the same with esters mixtures. However, we could still see that there were some differences in the details of their contour maps. For instance, the antagonism degree of mixture BA+EB (Figure 2d) was weaker when the $\ln\text{OAV}_{\text{BA}}$ and $\ln\text{OAV}_{\text{EB}}$ values were close to 2–3. So did the mixture EA+BA when the $\ln\text{OAV}_{\text{EA}}$ and $\ln\text{OAV}_{\text{BA}}$ values were close to 4.5–5.5 (Figure 2b). A very obvious difference was that there were almost no olfactory measured sample data in the areas mentioned above. Because there were not enough samples to reflect the real odor interaction in these areas, the performance of the SVR model to the corresponding area was easily affected by other samples. In machine learning, this phenomenon is generally observed because of insufficient sample amount and lacking data representativeness [42]. The samples in the

aldehyde mixtures were more evenly dispersed, so the odor interaction in each area was fully reflected. Therefore, reasonable sampling also should be concerned when using machine learning methods in odor researches.

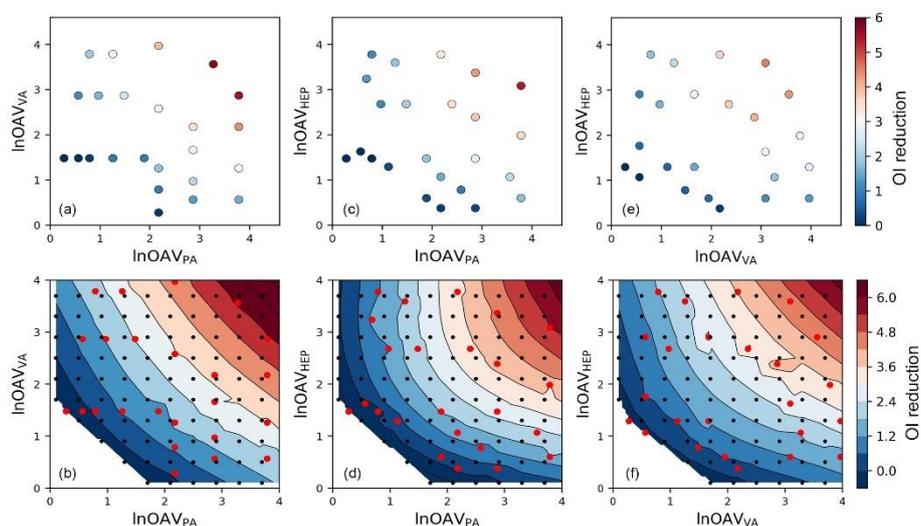


Figure 3. The scatter plot of the relationship between olfactory measured odor intensity reduction degree (*OI reduction*) and components' *lnOAV* values, and corresponding contour map after adding SVR model predicted data (in the contour map, red dots: olfactory measured data; black dots: SVR predicted data) for binary mixture (a,b) PA+VA, (c,d) PA+HEP, and (e,f) VA+HEP.

The contour maps of binary aromatic hydrocarbon mixtures were plotted in Figure 4. Unlike binary mixtures of esters and aldehydes, the overall antagonism degree of binary aromatic hydrocarbon mixtures was at a relatively low level. When the *lnOAV* value of both components was higher than 2.5, a similar antagonism degree-mixing ratio relationship like esters and aldehydes was observed (Figure 4b,d,f). When their *lnOAV* values were smaller than this critical value, we observed a synergism effect (i.e., the negative *OI reduction* value which meant that the OI_{mix} was higher than the OI_{sum}). Because there was no actual olfactory measured sample data in this area, the reliability of this phenomenon should be verified by further olfactory evaluation tests. In all these contour maps, the odor interaction of odor samples with too small *lnOAV* values (i.e., the lower left blank corner of the contour plot) was not considered. Because odor samples belonging to this area usually had the odor intensity value lower than 2.0 of the OIRS (it could be demonstrated from Figure 1). For those odor samples, the error of olfactory evaluation was usually higher [25]. Meanwhile, it also was more meaningful to analyze the odor interaction of odor mixtures with distinct olfactory stimulation.

3.3. Similarity of Binary Odor Interaction Pattern

In order to further verify the above-observed odor interaction pattern, we also analyzed the influence of the sample's odor intensity level. As depicted in Figure 5, all the olfactory measured data and SVR predicted data were employed and colors represented the odor intensity of each sample. We defined the *OI reduction ratio* (Equation (7)) and mixing ratio of the binary mixture (i.e., $x_a = \ln OAV_a / (\ln OAV_a + \ln OAV_b)$). Firstly, we observed the same conclusion as the above contour maps. When the *lnOAV* mixing ratio of the two components was close, the antagonism degree was the most obvious (i.e., higher *OI reduction ratio*). Secondly, most of the odor samples followed the same odor interaction pattern regardless of its specific odor intensity level. No significant correlation was observed between the *OI reduction ratio* and the sample's odor intensity value. It was consistent with the phenomenon observed in our previous PDE (partial differential equation) model researches [25]. Compared with the influence of the sample's odor intensity value, the odor interaction degree was mainly affected by the components' mixing ratio.

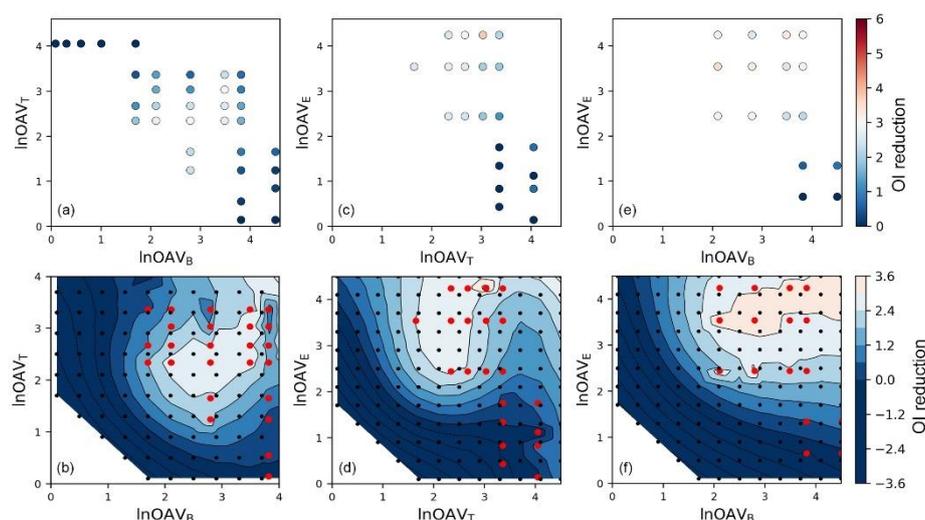


Figure 4. The scatter plot of the relationship between olfactory measured odor intensity reduction degree (*OI reduction*) and components' $\ln OAV$ values, and corresponding contour map after adding SVR model predicted data (in the contour map, red dots: olfactory measured data; black dots: SVR predicted data) for binary mixture (a,b) B+T, (c,d) T+E, and (e,f) B+E.

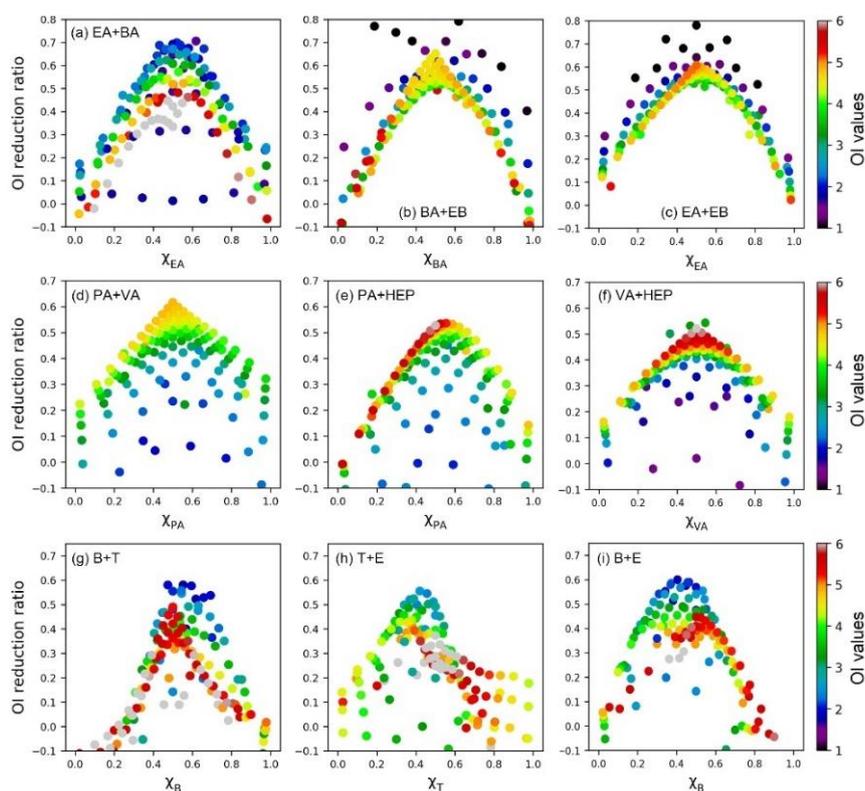


Figure 5. The scatter plot of *OI reduction ratio* vs. components' mixing ratio for binary odor mixtures. Colors represent the odor intensity of each sample.

In this study, we mainly employed the SVR algorithm as a useful tool for data analysis. Based on its strong regression ability, more reliable data was collected and it helped to explore the odor interaction more intuitively. Although the odor interaction has been widely studied by many empirical models who have made very accurate explanations [5,6,11,43], the machine learning method still has distinct advantages like visual analysis and low time/economic cost. As we found in this study, enough olfactory measured data was an essential guarantee to the accuracy of machine learning models. When applying

the machine learning methods, we also should pay attention to the sample representativeness and the objective analysis of simulation results. On the other hand, the observed odor interaction pattern was only verified by several substances from the esters, aldehydes, and aromatic hydrocarbon groups. In order to further prove the reliability and applicable scope of currently observed odor interaction pattern, it is necessary to test more odor substances.

4. Conclusions

In this study, a support vector regression algorithm was employed to train an odor intensity prediction model of binary esters, aldehydes, and aromatic hydrocarbon mixtures individually. The chemical concentrations of each component were directly transformed into the odor intensity of the mixture, and it successfully avoided the interference from odor threshold measurement and individual components' odor intensities evaluation which was usually performed by human assessors. The optimized model showed high accuracy for both the training samples and test samples. It was also considered adequate for other odor mixtures with different mixing ratios and concentration levels. Based on the support vector regression model, more odor data were collected and these data supported the visual analysis of odor interaction. Compared with traditional empirical models, the visual analysis method was more intuitive and provided more information. A similar odor interaction pattern was observed among these binary odor mixtures. Meanwhile, the importance of original olfactory measured data and data representativeness was also proved in the results. As a fast-growing technology, we have demonstrated its potential in the odor interaction analysis.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1424-8220/20/6/1707/s1>, Code S1: Code of SVR model for odor interaction simulation.ipynb.

Author Contributions: Conceptualization, L.Y. and J.L.; methodology, C.W.; visualization, software and writing, L.Y.; supervision, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key R&D Program of China (No. 2016YFC0700600; No. 2017YFB0702100), and the National Natural Science Foundation of China (No. 51601014).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jiang, G.M.; Melder, D.; Keller, J.; Yuan, Z.G. Odor emissions from domestic wastewater: A review. *Crit. Rev. Environ. Sci. Technol.* **2017**, *47*, 1581–1611. [[CrossRef](#)]
2. Le Berre, E.; Beno, N.; Ishii, A.; Chabanet, C.; Etievant, P.; Thomas-Danguin, T. Just noticeable differences in component concentrations modify the odor quality of a blending mixture. *Chem. Senses* **2008**, *33*, 389–395. [[CrossRef](#)] [[PubMed](#)]
3. Lewkowska, P.; Dymerski, T.; Gebicki, J.; Namiesnik, J. The use of sensory analysis techniques to assess the quality of indoor air. *Crit. Rev. Anal. Chem.* **2017**, *47*, 37–50. [[CrossRef](#)] [[PubMed](#)]
4. Atanasova, B.; Langlois, D.; Nicklaus, S.; Chabanet, C.; Etievant, P. Evaluation of olfactory intensity: Comparative study of two methods. *J. Sens. Stud.* **2004**, *19*, 307–326. [[CrossRef](#)]
5. Wu, C.D.; Liu, J.M.; Zhao, P.; Piringner, M.; Schaubberger, G. Conversion of the chemical concentration of odorous mixtures into odour concentration and odour intensity: A comparison of methods. *Atmos. Environ.* **2016**, *127*, 283–292. [[CrossRef](#)]
6. Wakayama, H.; Sakasai, M.; Yoshikawa, K.; Inoue, M. Method for predicting odor intensity of perfumery raw materials using dose-response curve database. *Ind. Eng. Chem. Res.* **2019**, *58*, 15036–15044. [[CrossRef](#)]
7. Sarkar, U.; Hobbs, S.E. Odour from municipal solid waste (MSW) landfills: A study on the analysis of perception. *Environ. Int.* **2002**, *27*, 655–662. [[CrossRef](#)]
8. Yan, L.C.; Liu, J.M.; Wang, G.H.; Wu, C.D. An odor interaction model of binary odorant mixtures by a partial differential equation method. *Sensors* **2014**, *14*, 12256–12270. [[CrossRef](#)]
9. Cain, W.S.; Schiet, F.T.; Olsson, M.J.; deWijk, R.A. Comparison of models of odor interaction. *Chem. Senses* **1995**, *20*, 625–637. [[CrossRef](#)]

10. Olsson, M.J. An integrated model of intensity and quality of odor mixtures. *Ann. N. Y. Acad. Sci.* **1998**, *855*, 837–840. [[CrossRef](#)]
11. Teixeira, M.A.; Rodriguez, O.; Rodrigues, A.E. The perception of fragrance mixtures: A comparison of odor intensity models. *AIChE J.* **2010**, *56*, 1090–1106. [[CrossRef](#)]
12. Yu, Z.M.; Guo, H.Q.; Lague, C. Development of a livestock odor dispersion model: Part II. Evaluation and validation. *J. Air Waste Manag.* **2011**, *61*, 277–284. [[CrossRef](#)] [[PubMed](#)]
13. Schiffman, S.S.; McLaughlin, B.; Katul, G.G.; Nagle, H.T. Eulerian-Lagrangian model for predicting odor dispersion using instrumental and human measurements. *Sens. Actuat. B Chem.* **2005**, *106*, 122–127. [[CrossRef](#)]
14. Olsson, M.J. An interaction-model for odor quality and intensity. *Percept. Psychophys.* **1994**, *55*, 363–372. [[CrossRef](#)] [[PubMed](#)]
15. Teixeira, M.A.; Rodriguez, O.; Rodrigues, A.E. Prediction model for the odor intensity of fragrance mixtures: A valuable tool for perfumed product design. *Ind. Eng. Chem. Res.* **2013**, *52*, 963–971. [[CrossRef](#)]
16. Butler, K.T.; Davies, D.W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine learning for molecular and materials science. *Nature* **2018**, *559*, 547–555. [[CrossRef](#)]
17. Vu, T.V.; Shi, Z.B.; Cheng, J.; Zhang, Q.; He, K.B.; Wang, S.X.; Harrison, R.M. Assessing the impact of clean air action on air quality trends in Beijing using a machine learning technique. *Atmos. Chem. Phys.* **2019**, *19*, 11303–11314. [[CrossRef](#)]
18. Arabgol, R.; Sartaj, M.; Asghari, K. Predicting nitrate concentration and its spatial distribution in groundwater resources using support vector machines (SVMs) model. *Environ. Model. Assess.* **2016**, *21*, 71–82. [[CrossRef](#)]
19. Vega García, M.; Aznarte, J.L. Shapley additive explanations for NO₂ forecasting. *Ecol. Inform.* **2020**, *56*, 101039. [[CrossRef](#)]
20. Vanarse, A.; Osseiran, A.; Rassau, A. Real-time classification of multivariate olfaction data using spiking neural networks. *Sensors* **2019**, *19*, 1841. [[CrossRef](#)]
21. Szulczynski, B.; Gebicki, J. Determination of odor intensity of binary gas mixtures using perceptual models and an electronic nose combined with fuzzy logic. *Sensors* **2019**, *19*, 3473. [[CrossRef](#)]
22. Zhu, L.T.; Jia, H.L.; Chen, Y.B.; Wang, Q.; Li, M.W.; Huang, D.Y.; Bai, Y.L. A novel method for soil organic matter determination by using an artificial olfactory system. *Sensors* **2019**, *19*, 3417. [[CrossRef](#)] [[PubMed](#)]
23. Yan, L.C.; Liu, J.M.; Jiang, S.; Wu, C.D.; Gao, K.W. The regular interaction pattern among odorants of the same type and its application in odor intensity assessment. *Sensors* **2017**, *17*, 1624. [[CrossRef](#)]
24. Yan, L.C.; Liu, J.M.; Fang, D. Use of a modified vector model for odor intensity prediction of odorant mixtures. *Sensors* **2015**, *15*, 5697–5709. [[CrossRef](#)]
25. Yan, L.C.; Liu, J.M.; Qu, C.; Gu, X.Y.; Zhao, X. Research on odor interaction between aldehyde compounds via a partial differential equation (PDE) model. *Sensors* **2015**, *15*, 2888–2901.
26. Krajewski, S.; Nowacki, J. Dual-phase steels microstructure and properties consideration based on artificial intelligence techniques. *Arch. Civ. Mech. Eng.* **2014**, *14*, 278–286. [[CrossRef](#)]
27. Mariette, A.; Rahul, K. *Efficient Learning Machines*, 1st ed.; Apress: New York, NY, USA, 2015; p. 67.
28. Wen, Y.F.; Cai, C.Z.; Liu, X.H.; Pei, J.F.; Zhu, X.J.; Xiao, T.T. Corrosion rate prediction of 3C steel under different seawater environment by using support vector regression. *Corros. Sci.* **2009**, *51*, 349–355. [[CrossRef](#)]
29. Azimi, H.; Bonakdari, H.; Ebtehaj, I. Design of radial basis function-based support vector regression in predicting the discharge coefficient of a side weir in a trapezoidal channel. *Appl. Water Sci.* **2019**, *9*, 78. [[CrossRef](#)]
30. Shi, H.; Xiao, H.; Zhou, J.; Li, N.; Zhou, H. Radial basis function kernel parameter optimization algorithm in support vector machine based on segmented dichotomy. In Proceedings of the 5th International Conference on Systems and Informatics (ICSAI), Nanjing, China, 10–12 November 2018; pp. 383–388.
31. Laref, R.; Losson, E.; Sava, A.; Siadat, M. Support vector machine regression for calibration transfer between electronic noses dedicated to air pollution monitoring. *Sensors* **2018**, *18*, 3716. [[CrossRef](#)] [[PubMed](#)]
32. Shin, D.; Yamamoto, Y.; Brady, M.P.; Lee, S.; Haynes, J.A. Modern data analytics approach to predict creep of high-temperature alloys. *Acta Mater.* **2019**, *168*, 321–330. [[CrossRef](#)]
33. Sun, Y.T.; Bai, H.Y.; Li, M.Z.; Wang, W.H. Machine learning approach for prediction and understanding of glass-forming ability. *J. Phys. Chem. Lett.* **2017**, *8*, 3434–3439. [[CrossRef](#)] [[PubMed](#)]

34. Kerckhoffs, J.; Hoek, G.; Portengen, L.; Brunekreef, B.; Vermeulen, R.C.H. Performance of prediction algorithms for modeling outdoor air pollution spatial surfaces. *Environ. Sci. Technol.* **2019**, *53*, 1413–1421. [[CrossRef](#)] [[PubMed](#)]
35. Tsirikoglou, P.; Abraham, S.; Contino, F.; Lacor, C.; Ghorbaniasl, G. A hyperparameters selection technique for support vector regression models. *Appl. Soft. Comput.* **2017**, *61*, 139–148. [[CrossRef](#)]
36. Brattoli, M.; de Gennaro, G.; de Pinto, V.; Loiotile, A.D.; Lovascio, S.; Penza, M. Odour detection methods: Olfactometry and chemical sensors. *Sensors* **2011**, *11*, 5290–5322. [[CrossRef](#)] [[PubMed](#)]
37. Deshmukh, S.; Jana, A.; Bhattacharyya, N.; Bandyopadhyay, R.; Pandey, R.A. Quantitative determination of pulp and paper industry emissions and associated odor intensity in methyl mercaptan equivalent using electronic nose. *Atmos. Environ.* **2014**, *82*, 401–409. [[CrossRef](#)]
38. Liu, T.P.; Zhang, W.T.; McLean, P.; Ueland, M.; Forbes, S.L.; Su, S.W. Electronic nose-based odor classification using genetic algorithms and fuzzy support vector machines. *Int. J. Fuzzy Syst.* **2018**, *20*, 1309–1320. [[CrossRef](#)]
39. Pan, L.; Yang, S.X. An electronic nose network system for online monitoring of livestock farm odors. *IEEE/ASME Trans. Mechatron.* **2009**, *14*, 371–376. [[CrossRef](#)]
40. Niu, Y.; Wang, P.; Xiao, Z.; Zhu, J.; Sun, X.; Wang, R. Evaluation of the perceptual interaction among ester aroma compounds in cherry wines by GC–MS, GC–O, odor threshold and sensory analysis: An insight at the molecular level. *Food Chem.* **2019**, *275*, 143–153. [[CrossRef](#)]
41. Wu, Q.; Wang, Z.; Hu, X.; Zheng, T.; Yang, Z.; He, F.; Li, J.; Wang, J. Uncovering the eutectics design by machine learning in the Al–Co–Cr–Fe–Ni high entropy system. *Acta Mater.* **2020**, *182*, 278–286. [[CrossRef](#)]
42. Chen, Z.Y.; Zhang, R.; Zhang, T.H.; Ou, C.Q.; Guo, Y. A kriging-calibrated machine learning method for estimating daily ground-level NO₂ in mainland China. *Sci. Total Environ.* **2019**, *690*, 556–564. [[CrossRef](#)]
43. Kim, K.H. Experimental demonstration of masking phenomena between competing odorants via an air dilution sensory test. *Sensors* **2010**, *10*, 7287–7302. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).