



Article

Catalytic Performance of Cycloalkyl-Fused Aryliminopyridyl Nickel Complexes toward Ethylene Polymerization by QSPR Modeling

Md Mostakim Meraz ^{1,2,†}, Arfa Abrar Malik ^{1,2,†}, Wenhong Yang ^{1,2,*}  and Wen-Hua Sun ^{1,2,*} 

¹ Key Laboratory of Engineering Plastics and Beijing National Laboratory for Molecular Science, Institute of Chemistry, Chinese Academy of Sciences, Beijing 100190, China; meraz@iccas.ac.cn (M.M.M.); arfmalik@iccas.ac.cn (A.A.M.)

² University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: whyang@iccas.ac.cn (W.Y.); whsun@iccas.ac.cn (W.-H.S.); Tel.: +86-10-62557955 (W.Y.); +86-10-62557955 (W.-H.S.)

† Md Mostakim Meraz and Arfa Abrar Malik contributed equally to this work.

Abstract: Quantitative structure–property relationship (QSPR) modeling is performed to investigate the role of cycloalkyl-fused rings on the catalytic performance of 46 aryliminopyridyl nickel precatalysts. The catalytic activities for nickel complexes in ethylene polymerization are well-predicted by the obtained 2D-QSPR model, exploring the main contribution from the charge distribution of negatively charged atoms. Comparatively, 3D-QSPR models show better predictive and validation capabilities than that of 2D-QSPR for both catalytic activity (*Act.*) and the molecular weight of the product (*M_w*). Three-dimensional contour maps illustrate the predominant effect of a steric field on both catalytic properties; smaller sizes of cycloalkyl-fused rings are favorable to *Act.*, whereas they are unfavorable to *M_w*. This study may provide assistance in the design of a new nickel complex with high catalytic performance.

Keywords: catalytic activity; cycloalkyl-fused derivatives; ethylene polymerization; nickel complexes; quantitative structure–property relationship



Citation: Meraz, M.M.; Malik, A.A.; Yang, W.; Sun, W.-H. Catalytic Performance of Cycloalkyl-Fused Aryliminopyridyl Nickel Complexes toward Ethylene Polymerization by QSPR Modeling. *Catalysts* **2021**, *11*, 920. <https://doi.org/10.3390/catal11080920>

Academic Editors: Christoffer Heath Turner, Tibor Szilvási, Wei An and Yaqiong Su

Received: 1 July 2021

Accepted: 28 July 2021

Published: 29 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Since the pioneering report of α -diimino-nickel precatalysts toward ethylene oligo-/polymerization in the middle of 1990s [1], numerous developments have been witnessed in the design and synthesis of nickel complex catalysts, showing their impact on a wide range of catalytic properties, such as catalyst activity, molecular weight, and the distribution of the produced oligomer/polymer [2,3]. In recent years, our group and others reported aryliminopyridyl nickel precatalysts with the incorporation of cycloalkyl-fused rings as compatible N,N-ligand frames. The fused rings rang from small sizes, like cyclopentyl and cyclohexyl, to bigger sizes, such as cycloheptyl and cyclooctyl, as in Scheme 1. All of these derivatives exhibit good activities in ethylene oligo-/polymerization and produce polyethylenes of generally low molecular weight [4,5]. Their specific properties show dependence on the fused ring size and the electronic and steric effects of the substituents. Even though the empirical information on variation trends can be obtained and provide useful guidance for experiments, it is more meaningful to explain the principle underlying the improved catalytic performance through a modeling method at the molecular level.

Category: A							Category: B							Category: C							Category: D							Category: E						
No.	Cat.	R ¹	R ²	R ³	X	Act. ^a	M _w ^b	No.	Cat.	R ¹	R ²	R ³	R	X	Act. ^a	M _w ^b	No.	Cat.	R ¹	R ²	X	Act. ^a	M _w ^b											
1	A	Me	Me	H	Br	3.44	2700	17	C	Me	Me	H	Cl	Cl	3.7	59	33	D	Me	H	Cl	0.30	84											
2	A	Et	Et	H	Br	1.13	4900	18	C	Et	Et	H	Cl	Cl	7.1	60	34	D	Et	H	Cl	0.29	86											
3	A	iPr	iPr	H	Br	0.60	3600	19	C	iPr	iPr	H	Cl	Cl	8.7	59	35	D	iPr	H	Cl	0.23	82											
4	A	Me	Me	Me	Br	2.18	3800	20	C	Me	Me	Me	Cl	Cl	2.4	62	36	D	Me	Me	Cl	0.34	82											
5	A	Et	Et	Me	Br	1.00	5100	21	C	Et	Et	Me	Cl	Cl	2.5	59	37	D	Et	Me	Cl	0.31	83											
6	B	Me	Me	H	Cl	7.18	2453	22	C	Me	Me	H	Ph	Cl	3.1	60	38	D	Me	H	Br	0.48	86											
7	B	Et	Et	H	Cl	6.22	3481	23	C	Et	Et	H	Ph	Cl	3.4	60	39	D	Et	H	Br	0.35	86											
8	B	iPr	iPr	H	Cl	4.71	4180	24	C	iPr	iPr	H	Ph	Cl	4.1	60	40	D	iPr	H	Br	0.32	84											
9	B	Me	Me	Me	Cl	6.85	2825	25	C	Me	Me	Me	Ph	Cl	2.5	62	41	D	Me	Me	Br	0.51	86											
10	B	Et	Et	Me	Cl	5.65	3830	26	C	Et	Et	Me	Ph	Cl	4.0	59	42	D	Et	Me	Br	0.37	86											
11	B	C ₆	C ₆	H	Br	3.08	2360	27	C	C ₆	C ₆	H	H	Br	4.85	1840	43	E	Me	H	Br	3.08	670											
12	B	C ₅	C ₅	H	Br	4.97	3640	28	C	C ₅	C ₅	H	H	Br	2.40	3210	44	E	Et	H	Br	1.01	750											
13	B	C ₆	Me	Me	Br	6.27	1910	29	C	C ₆	Me	Me	H	Br	4.03	1720	45	E	Me	Me	Br	3.51	870											
14	B	C ₆	C ₆	H	Cl	3.47	2320	30	C	C ₆	Me	Me	H	Cl	3.39	2190	46	E	Et	Me	Br	1.45	730											
15	B	C ₅	C ₅	H	Cl	4.76	3210	31	C	Me	Me	Me	H	Br	4.10	2240																		
16	B	C ₆	Me	Me	Cl	4.17	1830	32	C	Me	Me	Me	H	Cl	2.98	1980																		

Scheme 1. Structures of the initial data set for 46 nickel complex precatalysts along with their catalytic activities and molecular weight of produced polyethylene under the optimum conditions. ^a, 10⁶ g·mol^{−1}·h^{−1}, ^b, g·mol^{−1}.

Quantitative structure–activity/property relationship (QSA/PR) modeling is an approach to build the quantitative relationship between the structure of a catalyst and its activity and/or property, through the key structural descriptors within the catalyst. This method has been mainly proposed for the field of drug design and has already proven a powerful tool. However, its application in the polyolefin catalyst is relatively limited. Regarding the field of polyolefin catalysts, QSPR has been performed to predict the catalytic activities and molecular weight of polyethylene by the metallocene complexes [6,7], bis(imino)pyridine iron complex systems [8], and aryliminopyridyl metal complexes [9,10]. Furthermore, by using the chemically meaningful descriptors, such as buried volume, the free energy of reactions are predicted for *ansa*-metallocenes catalysts towards propene polymerization and ethene/1-hexene copolymerization [11–14]. Recently, QSPR modeling was applied to investigate the role of cycloalkyl-fused rings within bis(imino)pyridine (**Fe**, **Co**, **Cr**) complexes on their catalytic performance [15]. The obtained results indicate the dominant role played by the conjugated degree in complexes on the catalytic activities.

However, this finding is not applicable for nickel complexes. As shown in Scheme 1, the introduction of a benzene ring in the aryliminopyridyl framework (category D) does not enhance the catalytic activities; on the contrary, their activities are lower by about ten times than that the analogues without a benzene ring (category C). It means that, although the ligands of the complexes are similar, the effects of a cycloalkyl-fused ring on the catalytic performance are varied, affected by the change of different central metal atoms and the framework of ligands. Therefore, it is necessary to further investigate the nickel system to detect the role of the cycloalkyl-fused ring on the corresponding catalytic performance. In this work, we subsequently apply both 2D- and 3D-QSPR methods to the

nickel analogue complexes to explore the variation mechanism caused by the different sizes of cycloalkyl-fused rings. 2D-QSPR modeling obtains the predictive model for catalytic activities. Through the analysis of selected descriptors, it was found that the charge distribution of net charged atom plays the dominant role, which is different from the results of cycloalkyl-fused aryliminopyridyl Fe/Co analogues, where the dominant role is played by the number of aromatic bonds. Meanwhile, the 3D-QSPR modeling reveals the important effect of steric fields in the models of both catalytic activity and molecular weight, which is in agreement with previous works [10,15]. An exception to this agreement is that the favorable areas are different. It indicates that, depending on the analogue complexes with different central metal atoms, the cycloalkyl-fused ring play different roles in the catalytic performance.

2. Results and Discussions

2.1. 2D-QSPR Modeling

In order to build the quantitative relationship between the structure of a complex and its catalytic performance, the descriptors are first calculated and selected for the model. Then, four parameters including correlation coefficient (R^2), the root mean square error of the fitted data ($RMSEF$), the cross-validation coefficient (Q^2), and the root mean squared error of cross-validation ($RMSEV$) are calculated and listed Table 1 to validate the stability and predictive power of the model. Additionally, the value of A means the number of components. As shown in Table 1, based on the VIP-PLS analysis, the number of descriptors gradually decreased from 63 to 22 and 12 to 5. An acceptable 2D-QSPR model should have higher R^2 and Q^2 values, as well as lower values of $RMSEV/RMSEF$ under a smaller number of descriptors. Therefore, a set of 12 descriptors is selected as the optimum number of independent variables.

Table 1. The values of the R^2 , Q^2 , $RMSEF$, and $RMSEV$ for the 2D-QSAR model for catalytic activity at different numbers of descriptors.

No.	A	R^2	Q^2 (k = 5)	$RMSEF$	$RMSEV$
63	6	0.9211	0.5810	0.6390	0.2772
22	5	0.8770	0.6892	0.5503	0.3461
12	5	0.8755	0.7328	0.5103	0.3483
5	2	0.7363	0.6731	0.5644	0.5069

The triangle matrix is plotted to show the correlation between each pair of descriptors, indicating their independent nature as in Figure 1. The specific values of each descriptor are summarized in Table S1 for each Ni complex.

To build and validate the QSPR model, the training and test sets are selected from the whole data by the Kennard-Stone method [16]. As given in Table S2, the combination of 29 complexes as a training set and 10 complexes as a test set is regarded as optimum. The correlation coefficients (R^2) for the training set and test set are 0.892 and 0.814, respectively. The cross-validation coefficient (Q^2) is 0.697, indicating the acceptable prediction and validation capabilities of the model. Comparatively, the R^2 values are not as high as in previous reports [9,10,15]. The reason might be the selection of model complexes, which not only vary in the sizes of the cycloalkyl-fused rings, but also vary in the presence of alkyl substituents on the fused ring, such as categories D and E in Scheme 1. The comparisons of experimental and predicted catalytic activities are plotted in Figure 2. The specific values of predicted catalytic activities for each complex are given in Table S3.

	No. 1																							
No. 1	1.00	No. 2																						
No. 2	0.46	1.00	No. 3																					
No. 3	0.49	0.33	1.00	No. 4																				
No. 4	0.29	0.23	0.80	1.00	No. 5																			
No. 5	0.40	0.52	0.61	0.50	1.00	No. 6																		
No. 6	0.01	0.00	0.10	0.25	0.05	1.00	No. 7																	
No. 7	0.34	0.48	0.25	0.17	0.61	0.00	1.00	No. 8																
No. 8	0.62	0.59	0.52	0.28	0.73	0.00	0.79	1.00	No. 9															
No. 9	0.15	0.51	0.25	0.17	0.37	0.00	0.37	0.40	1.00	No. 10														
No. 10	0.64	0.59	0.24	0.11	0.26	0.02	0.32	0.44	0.19	1.00	No. 11													
No. 11	0.26	0.15	0.59	0.54	0.55	0.17	0.12	0.25	0.07	0.11	1.00	No. 12												
No. 12	0.12	0.23	0.22	0.33	0.33	0.13	0.35	0.19	0.13	0.20	0.32	1.00	Act.											
Act.	0.38	0.60	0.56	0.45	0.55	0.11	0.45	0.53	0.25	0.37	0.26	0.34	1.00											

Figure 1. The correlations matrix for the selected 12 descriptors along with catalytic activity.

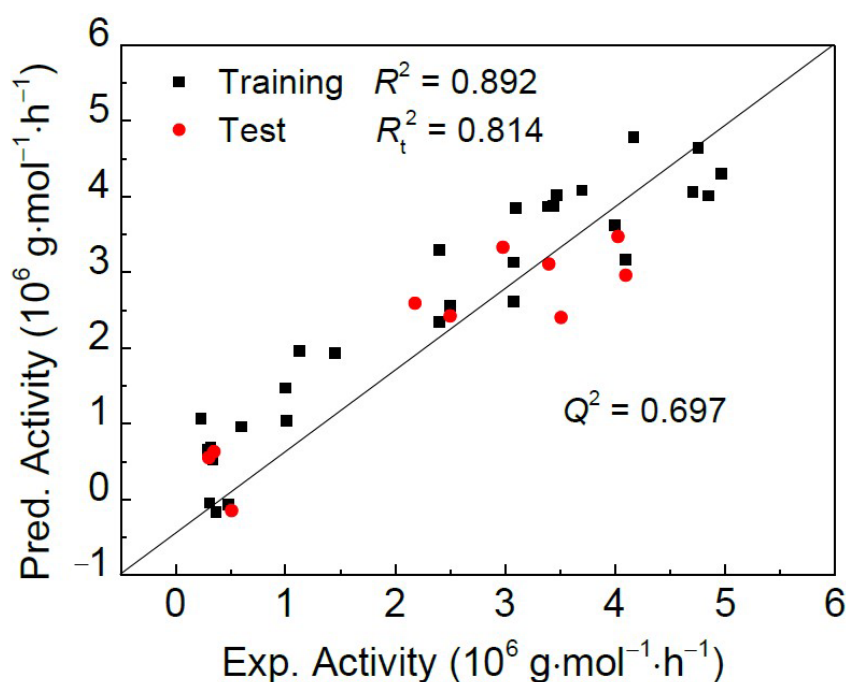


Figure 2. Experimental catalytic activities versus predicted values by the 2D-QSPR model for the training set and test set.

In order to interpret the obtained model, detailed information on the selected 12 descriptors is listed in Table 2. Besides the R^2 values of each descriptor in the fitting model, the contribution values to catalytic activities are also calculated and listed together. The descriptors were ranked from higher to lower according to their contribution values to catalytic activity. The explanations focus on the descriptors exhibiting higher contributions. It is clear that among all the selected descriptors, the minimum partial charges on a C atom (No.2), present the biggest contribution to the catalytic activities, with the value of 29.17%. This electrostatic descriptor describes the characteristics of the charge distribution of C atoms, showing that the minimum charge values on carbon atoms present largely positive correlations with catalytic activity.

Table 2. Regression coefficients and the contribution of each molecular descriptor in the final model.

No.	Molecular Descriptor	Coeff.	Contr.(%)
2	Min partial charge for a C atom [Zefirov's PC]	0.7854	29.17
3	WNSA-3 Weighted PNSA (PNSA3*TMSA/1000) [Zefirov's PC]	0.6168	17.38
9	HA dependent HDCA-2/TMSA [Quantum-Chemical PC]	−0.3804	12.74
12	Min valency of a H atom	−0.2534	9.21
11	Avg valency of a N atom	0.4056	9.15
10	Max SIGMA-SIGMA bond order	0.1834	5.47
8	WNSA-3 Weighted PNSA (PNSA3*TMSA/1000)[Quantum-Chemical PC]	0.1593	5.32
1	Average Information content (order 2)	0.1381	3.85
6	Max 1-electron react. index for a C atom	0.1501	3.59
5	Max electroph. react. index for a C atom	0.0899	3.29
7	FNSA-2 Fractional PNSA(PNSA-2/TMSA)[Quantum-Chemical PC]	−0.0123	0.44
4	HOMO–LUMO energy gap	−0.0141	0.40

Following the highest contribution to activity, the second and third highest contributors are descriptors No.3 and No.9, with the values of 17.38% and 12.74%, respectively. These two descriptors are classified as the charged partial surface areas (CPSA) descriptors [17,18], in which the No.3 descriptor is calculated by Equation (1):

$$WNSA3 = \frac{PNSA3}{TMSA} = \left(\sum_a q_a^- \times SA_a^- \right) / TMSA \quad (1)$$

where q_a^- is the partial charge for the a th negatively charged atom, SA_a^- is the surface area contribution of this a th atom, and TMSA is the total molecular solvent-accessible surface area. For the model complexes, the negatively charged atoms usually refer to halogen, nitrogen, and part of the carbon atoms. Therefore, this descriptor roughly reflects the charge distribution of these negatively charged atoms. Similar to descriptor No.2, descriptor No.3 presents a positive contribution to catalytic activities as well, due to the fact that the minimum partial charge on the C atom is usually negative.

The No.9 descriptor is a parameter to describe the area-weighted surface charge of H-bonding donor H atoms, which is defined by Equation (2):

$$DCA2 = \left(\sum_D \frac{q_D \sqrt{S_D}}{\sqrt{S_{tot}}} \right), \quad D \in H_{H-donor} \quad (2)$$

where S_D is the solvent-accessible surface area of H-bonding donor H atoms, q_D is the partial charge on H-bonding donor H atoms, and S_{tot} is the total solvent-accessible molecular surface area. This descriptor shows the charge distribution on the H atoms and the interaction of hydrogen bonding. Different from the No.2 and No.3 descriptors, the No.9 descriptor has a negative correlation with catalytic activities, due to the generally positive charge of the H atoms.

These three electrostatic descriptors, reflecting the charge distribution on negatively charged atoms and H-bonding donor H atoms, provide around 60% of the contribution to catalytic activities, showing the dominant role of the net charge of negatively charged atoms on catalytic activities.

For the remaining descriptors, their explanations are briefly provided in the Supporting Materials.

2.2. 3D-QSPR Modeling

Different from the 2D-QSPR method, 3D-type builds the quantitative relationship based on the steric fields and electrostatic fields of ligands at the three-dimensional lattice. After modeling, the calculated values of the catalytic activity and molecular weight of the product are obtained and compared with the experimental data, as shown in Figure 3. The training and test sets are the same as those used in the 2D type. The detailed predicted

values of the catalytic activity and molecular weight of the product for each complex are given in Tables S4 and S5, respectively.

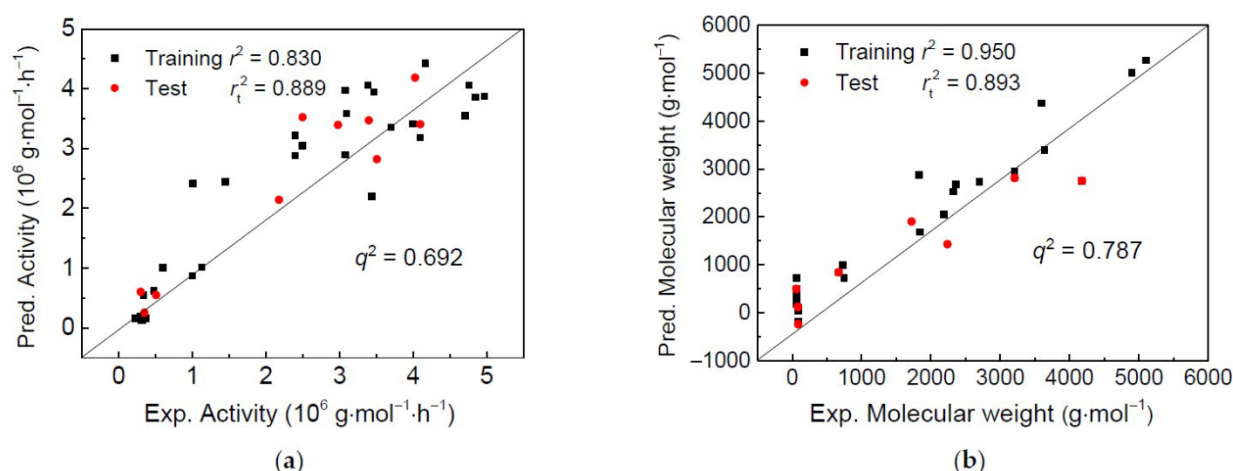


Figure 3. Experimental catalytic activities (a) and molecular weight (b) versus the predicted values by 3D-QSPR model for the training set and test set.

Regarding the model of catalytic activity, the correlation coefficient values for the training set (r^2) and test set (r_t^2) are 0.830 and 0.889, respectively, as shown in Figure 3a. Meanwhile, the cross-validation coefficient (q^2) value is 0.692, indicating the good predictive and validated capabilities of the model. The contributions for steric and electrostatic fields are calculated with values of 87.3% and 12.7%, respectively, indicating the predominant role of steric fields on catalytic activity. As to the model of molecular weight, quantitative results are obtained as in Figure 3b. The correlation coefficients (r^2 , r_t^2) for the training and test sets are 0.950 and 0.893, respectively, and the q^2 value is 0.787. Comparatively, the predictive and validated powers are higher than that of the model for catalytic activity. The decisive effect on molecular weight comes from the steric field, with an overwhelming contribution of 89.1%.

Since the dominant role is played by the steric effect in the models of both catalytic activity and the molecular weight of the product, only the 3D contour maps of steric fields are shown in Figure 4. The *iso*-surface values are the result of the multiplication of the standard deviation (StDev) of energy by the PLS coefficient (Coeff), providing an approximate location to explain the selective areas about its structure-performance relationship. Regarding the 3D contour map of the steric field for catalytic activity, as shown in Figure 4a, the green areas represent the positive effect of the bulky group, meaning the introduction of a bulky group in this region enhances the catalytic activity. Yellow areas, on the contrary, represent the negative effect of the bulky group. Clearly, we can see that the green areas are mainly located around the halogen atoms, and the yellow areas are surrounded on the upper side of the cycloalkyl-fused ring. This indicates that bigger, bulky steric substituents surrounding halogen atoms are favorable for catalytic activity, whereas a larger size of cycloalkyl-fused ring and/or bigger hindrance groups on the fused ring are unfavorable factors. As in Scheme 1, complexes in category D have propyl groups as substituents on the cyclohexane-fused ring and aryl-fused rings as substituents on pyridine; as a result, the catalytic activities for this category are lower than that of category C. This can be explained as: bigger steric groups on the frame work decrease the freedom degree of a complex and make its molecule rigid. As a result, the insertion reaction between the ethylene monomer and the central metal is hindered during coordinated polymerization, leading to the decrease in catalytic activity.

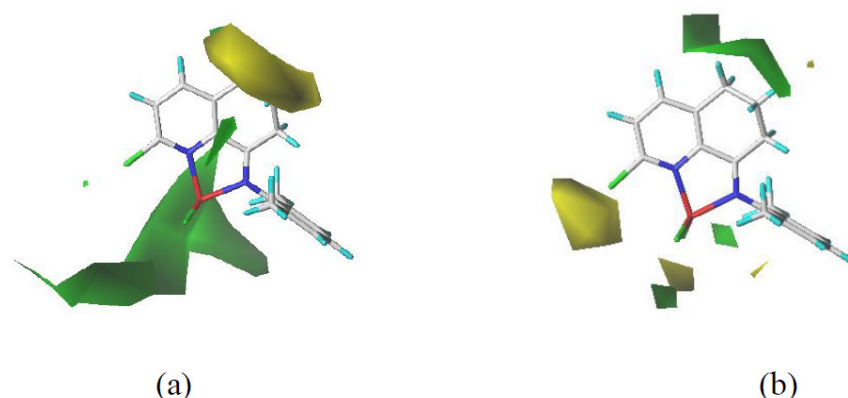


Figure 4. The 3D contour maps of the steric field for the model of (a) catalytic activity, and (b) molecular weight of product. The green and yellow regions indicate the positive and negative effects of the bulky group on catalytic performance, respectively.

The opposite of the results for the catalytic activities, the 3D contour map of steric fields for molecular weight show a completely different trend. It is clear in Figure 4b that green areas are located at the upper side of the cycloalkyl-fused ring, indicating that larger cycloalkyl-fused members and/or bulky substituents on the fused ring are favorable factors. This is because the highly crowded aniline group reduces the space and the rotation freedom surrounding the active site so that the monomer insertion can only occur from one side. Then, the β -H elimination is inhibited, resulting in highly linear polymers with a high molecular weight.

Comparing the 2D- and the 3D-QSPR, it is clear that the 3D type provides better predictive and validation capabilities for each catalytic property, due to the more exact and detailed information of spatial descriptors provided by the 3D type. Additionally, for the models of both catalytic activity and molecular weight, 3D results show that the steric field plays a dominant role. The favorable areas with respect to catalytic activity usually turn out to be the unfavorable areas for the molecular weight of the product.

3. Computational Methods

3.1. Data Set

A data set of 46 cycloalkyl-fused aryliminopyridyl nickel complexes is initially selected from previous experimental works [19–25]. As in Scheme 1, the structures of complexes include five categories with different membered fused rings within the framework. The 10-aryliminocycloocta[b]pyridylnickel bromides (category A) are synthesized from the reactions of Cycloocta[b]pyridin-10-one imino derivatives with (DME)NiBr₂ [19]. Regarding category B, the 9-(2-Cycloalkylphenylimino)-5,6,7,8-tetrahydrocycloheptapyridylnickel complexes are found from the condensation reactions of 5,6,7,8-tetrahydrocycloheptapyridine-9-one with various anilines [20,25]. As to category C, N-(2-substituted-5,6,7-trihydroquinolin-8-ylidene)-arylammoniumnickel(II) dichlorides (complex 17–26, Scheme 1) are synthesized by the one-pot stoichiometric reaction of nickel dichloride, 2-chloro- or 2-phenylsubstituted 5,6,7-trihydroquinolin-8-one, and the corresponding anilines [21]. Subsequently, the 8-(2-cycloalkylphenylimino)-5,6,7-trihydroquinoline derivatives react with (DME)NiBr₂ or NiCl₂ to form the corresponding cycloalkyl-substituted 8-arylimino-5,6,7-trihydroquinolylnickel halides (complexes 27–32, Scheme 1) [22]. Propyl substituted 4-arylimino-1,2,3-trihydroacridylnickel dihalide complexes (category D) are prepared by the metal-induced template reaction with NiCl₂·6H₂O or (DME)NiBr₂ [23]. As with category B, the category E complexes are synthesized from the condensation of 2-chloro-6,6-dimethylcyclopenta[b]pyridin-7-one with the corresponding aniline [24]. The observed catalytic activities and the molecular weight of the products are also listed in Scheme 1, which are taken under optimum conditions [15]. This is because the ethylene polymerization reaction actually takes place after the precatalyst is activated into the active species, and

optimum conditions guarantee that all the precatalysts can be activated. The detailed optimum conditions are listed in Table S6 for each complex. By Hotelling's T^2 method [26] in principle component analysis (PCA) [27], seven complexes (6, 7, 9, 10, 13, 18, and 19) are identified as outliers and removed. Further analyses are conducted on the rest of the 39 complexes.

3.2. QSPR Modeling

Based on the methodology discussed in previous reports [9,10,15], a brief introduction of 2D- and 3D-QSPR is described as follows. Regarding the 2D type, firstly, the geometry is optimized for each complex by density functional theory (DFT) calculations through the Dmol3 program [25]. GGA-type BP exchange-correlation functional [28–30] and double-numerical basis sets with polarization functions (DNP) basis sets [31,32] are chosen to calculate the geometry and energy combined with effective core potentials [33,34]. The comparisons of the optimized structures with the experimental crystal data are listed in Table S7, showing the reasonable optimized structures with lower standardized deviation values for the bond lengths and the bond angles. In order to use the CODESSA Software [35] to calculate the descriptors for each complex, the input file needs to be converted into a Gaussian type (*.log). Therefore, the single-point energy (SPE) calculation is carried out using the Gaussian 09 program package [36] under BLYP/6-31G* level based on a previous study [9]. The natural bond orbital (NBO) is used for electron distribution analysis [37,38].

The descriptors are then calculated by the CODESSA software, generating around 350 descriptors for each complex. As in previous studies [39–42], there are seven self-defined descriptors which show good correlations with catalytic activity. Therefore, these descriptors are also calculated for each complex as shown in Table S8. Adding these seven self-defined descriptors into the pool of descriptors, the total number of descriptors is around 357. This huge number of descriptors is first pre-screened by the heuristic method [43], leaving 99 descriptors for each complex. At this point, the values of the correlation coefficients (R^2) are calculated and show lower values below 0.6 for the model of the molecular weight of the product. Therefore, the modeling is proceeded only for the properties of catalytic activity. The obtained descriptors are cross-validated for the optimum values of the cross-validation coefficient (Q^2) and the root mean squared error of cross-validation (RMSEV) using the k-fold method [44]. Descriptors are then further reduced with a k-fold value of 5, by the variable importance of projection (VIP) [45,46] method incorporated in the partial least-squares (PLS) toolbox [47,48]. The average of the squared VIP scores equals 1, thus the “greater than one rule” is commonly used as a criterion for variable selection, meaning that only the descriptors with an importance value over 1 are selected. This pruning is repeated cyclically to gradually eliminate the descriptors.

To decide on the optimal number of descriptors, four parameters expressing the stability and predictive power of the model are employed. The correlation coefficient (R^2) is the square correlation coefficient between the experimental and predicted values of the activity, and the root mean square error of the fitted data (RMSEF) illustrates the difference between the predicted and experimental activities, which are defined as Equations (3) and (4), respectively.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

$$RMSEF = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4)$$

where \hat{y}_i and y_i are the predicted and experimental activity, respectively, \bar{y} is the averaged value of the experimental activity, and n is the number of complexes. The five-fold cross-

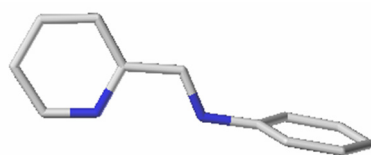
validation result is defined as the cross-validation coefficient (Q^2) in Equation (5), and the root mean squared error of cross-validation ($RMSEV$) is calculated by Equation (6):

$$Q^2 = \frac{\sum_{j=1}^4 \left(1 - \frac{\sum_{i=1}^n (y_v - \hat{y}_v)_i^2}{\sum_{i=1}^n (y_v - \bar{y}_v)_i^2} \right)}{4} \quad (5)$$

$$RMSEV = \sqrt{\frac{\sum_{i=1}^n (y_v - \hat{y}_v)_i^2}{n}} \quad (6)$$

where \hat{y}_v and y_v are the predicted and experimental activity in the validation set, respectively, \bar{y}_v is the averaged value of experimental activity for the complexes across all validation sets, and n is the number of complexes. A 2D-QSPR model is considered to be predictive, provided it has R^2 and Q^2 values close to 1.0, as well as values of $RMSEF/RMSEV$ close to 0. Later on, the data set is divided into a training set and a test set to build and validate the model, respectively.

3D-QSPR relates the catalytic properties of the complexes to their 3D descriptors, including steric and electrostatic types, which are calculated by comparative molecular field analysis (CoMFA) [49,50]. First, a database of all optimized structures is built, and according to the alignment rule, a common substructure is selected as the alignment structure, as shown in Scheme 2. After the alignment, structures are organized with the associated data into molecular spread sheets. A cube of dimension $20 \text{ \AA} \times 20 \text{ \AA} \times 20 \text{ \AA}$ is selected, in which molecular descriptors are calculated in a 3D grid at a step size of 2 \AA [49].



Scheme 2. The common framework selected from the complexes for the alignment of structures.

In the CoMFA method, the steric and electrostatic interactions between a probe atom, usually a carbon atom of sp^3 type corresponding to atom C.3 in the Tripos force field, and complex structures are calculated at each 3D grid point. As to the electrostatic field, coulomb interactions are calculated with the charge values assigned by the MMFF94 force field. Lennard-Jones potentials are used to assess steric fields by the interaction of van der Waals. The calculated 3D descriptors are not independent variables, thus PLS analysis has been performed to extract the value of the principle components. In order to investigate the predictive power of the models and to obtain the optimal number of principal components, leave-one-out (LOO) cross-validation is carried out to provide the result of the cross-validation coefficient (q^2). Afterwards, non-cross-validated analysis (without any validation) is performed to get the value of predictive correlation coefficient (r^2). Herein, q^2 and r^2 are described in Equations (7) and (8), respectively.

$$q^2 = 1 - \frac{\text{PRESS}}{\text{SD}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (7)$$

$$r^2 = \frac{\text{SD} - \text{PRESS}}{\text{SD}} = \frac{\left[\sum_{i=1}^n (y_i - \bar{y})^2 \right] - \left[\sum_{i=1}^n (y_i - \hat{y}_i)^2 \right]}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (8)$$

where \hat{y}_i , y_i , \bar{y} , and n are same as in Equation (4). A good 3D-QSPR model should have values of q^2 over 0.5 and r^2 over 0.6.

4. Conclusions

In the present work, QSPR modeling is performed to investigate the catalytic performance of cycloalkyl-fused aryliminopyridyl nickel precatalysts toward ethylene polymerization, including the catalytic activity and molecular weight of the product. The 2D-QSPR model shows acceptable predictive and validation capabilities for catalytic activity, indicating that the net charge of negatively charged atoms within complexes has a large positive contribution to catalytic activity. Meanwhile, 3D-QSPR analysis obtains good models not only for the catalytic activity, but also for the molecular weight of the product. A 3D contour map indicates the dominant role of steric fields for the both two performances. Larger steric environments around cycloalkyl-fused rings have an unfavorable effect on catalytic activity. On the contrary, these negative effect areas become positive effect areas in the model of the molecular weight of the product. This study explores the structural factors determining the catalytic performance of nickel precatalysts, providing potential assistance for the design of new nickel complexes with desirable properties.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/catal11080920/s1>, Table S1: The calculated values of the final 12 descriptors for each complex; Table S2: The modeling results of the different combinations of training and test sets including the R^2 , Q^2 , R_t^2 , $RMSEF$, and $RMSEV$ values for 39 complexes; Table S3: The values of the experimental and predicted catalytic activities for the training set of 29 complexes and the test set of 10 complexes from the 2D-QSAR analysis; Table S4: The values of the experimental and predicted catalytic activity for the training and test sets from the 3D-QSPR analysis; Table S5: The values of the experimental and predicted molecular weight for the training and test sets from the 3D-QSPR analysis; Table S6: Experimental values of catalytic activities, molecular weights, and melting temperatures of products under the optimum reaction conditions; Table S7: The calculated bond lengths and bond angles by DFT compared with experimental crystal structures for complex **8** along with the values of the standard deviation (δ); and Table S8: The values of different self-defined descriptors for each Ni complex.

Author Contributions: Conceptualization, W.Y.; methodology, M.M.M., W.Y.; software, A.A.M., W.Y.; validation, M.M.M., W.Y. and A.A.M.; formal analysis, M.M.M.; investigation, A.A.M.; resources, W.Y.; data curation, A.A.M.; writing—original draft preparation, W.Y. and M.M.M.; writing—review and editing, W.Y. and W.-H.S.; visualization, W.Y.; supervision, W.Y. and W.-H.S.; project administration, W.Y. and W.-H.S.; funding acquisition, W.Y. and W.-H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the Innovated Cultivation Project of ICCAS (CXPY-19) and the National Natural Science Foundation of China (No. 21871275).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Johnson, L.K.; Killian, C.M.; Brookhart, M. Pd(II)- and Ni(II)-Based Catalysts for Polymerization of Ethylene and α -Olefins. *J. Am. Chem. Soc.* **1995**, *23*, 6414–6415. [CrossRef]
2. Mecking, S. Cationic nickel and palladium complexes with bidentate ligands for the C–C linkage of olefins. *Coord. Chem. Rev.* **2000**, *203*, 325–351. [CrossRef]
3. Leung, D.H.; Ziller, J.W.; Guan, Z. Axial donating ligands: A new strategy for late transition metal olefin polymerization catalysis. *J. Am. Chem. Soc.* **2008**, *24*, 7538–7539. [CrossRef] [PubMed]
4. Bianchini, C.; Giambastiani, G.; Luconi, L.; Meli, A. Olefin oligomerization, homopolymerization and copolymerization by late transition metals supported by (imino) pyridine ligands. *Coord. Chem. Rev.* **2010**, *254*, 431–455. [CrossRef]
5. Wang, Z.; Liu, Q.; Solan, G.A.; Sun, W.-H. Recent advances in Ni-mediated ethylene chain growth: Nimine-donor ligand effects on catalytic activity, thermal stability and oligo-/polymer structure. *Coord. Chem. Rev.* **2017**, *350*, 68–83. [CrossRef]
6. Cruz, V.; Ramos, J.; Munoz-Escalona, A.; Lafuente, P.; Pena, B.; Martinez-Salazar, J. 3D-QSAR analysis of metallocene-based catalysts used in ethylene polymerisation. *Polymer* **2004**, *45*, 2061–2072. [CrossRef]

7. Martínez, S.; Cruz, V.L.; Ramos, J.; Martínez-Salazar, J. Polymerization activity prediction of zirconocene single-site catalysts using 3D quantitative structure–activity relationship modeling. *Organometallics* **2012**, *31*, 1673–1679. [\[CrossRef\]](#)
8. Cruz, V.L.; Martinez, J.; Martinez-Salazar, J.; Ramos, J.; Reyes, M.L.; Toro-Labbe, A.; Gutierrez-Oliva, S. QSAR model for ethylene polymerisation catalysed by supported bis (imino) pyridine iron complexes. *Polymer* **2007**, *48*, 7672–7678. [\[CrossRef\]](#)
9. Yang, W.; Yi, J.; Ma, Z.; Sun, W.-H. 2D-QSAR modeling on the catalytic activities of 2-azacycyl-6-aryliminopyridylmetal precatalysts in ethylene oligomerization. *Catal. Commun.* **2017**, *101*, 40–43. [\[CrossRef\]](#)
10. Yang, W.; Ma, Z.; Yi, J.; Ahmed, S.; Sun, W.-H. Catalytic performance of bis (imino) pyridine Fe/Co complexes toward ethylene polymerization by 2D-/3D-QSPR modeling. *J. Comput. Chem.* **2019**, *40*, 1374–1386. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Manz, T.A. Deactivation of Ti and Zr half-metallocene complexes activated with B(C₆F₅)₃: A case study in constructing DFT-based QSARs to predict unimolecular rate constants. *RSC Adv.* **2015**, *5*, 48246–48254. [\[CrossRef\]](#)
12. Falivene, L.; Cavallo, L.; Talarico, G. Buried volume analysis for propene polymerization catalysis promoted by group 4 metals: A tool for molecular mass prediction. *ACS Catal.* **2015**, *5*, 6815–6822. [\[CrossRef\]](#)
13. Ehm, C.; Vittoria, A.; Goryunov, G.P.; Izmer, V.V.; Kononovich, D.S.; Kulyabin, P.S.; Girolamo, R.D.; Budzelaar, P.H.M.; Voskoboynikov, A.Z.; Busico, V.; et al. A Systematic Study of the Temperature-Induced Performance Decline of ansa-Metallocenes for iPP. *Macromolecules* **2020**, *53*, 9325–9336. [\[CrossRef\]](#)
14. Ehm, C.; Vittoria, A.; Goryunov, G.P.; Izmer, V.V.; Kononovich, D.S.; Samsonov, O.V.; Budzelaar, P.H.M.; Voskoboynikov, A.Z.; Busico, V.; Uborsky, D.V.; et al. On the limits of tuning comonomer affinity of ‘Spaleck-type’ ansa-zirconocenes in ethene/1-hexene copolymerization: A high-throughput experimentation/QSAR approach. *Dalton Trans.* **2020**, *49*, 10162–10172. [\[CrossRef\]](#)
15. Yang, W.; Ahmed, S.; Fidelis, T.T.; Sun, W.-H. Effect of cycloalkyl-fused ring on the catalytic performance of bis (imino) pyridine metal complexes by QSPR modeling. *Catal. Commun.* **2019**, *132*, 105820–105824. [\[CrossRef\]](#)
16. Kennard, R.W.; Stone, L.A. Computer aided design of experiments. *Technometrics* **1969**, *11*, 137–148. [\[CrossRef\]](#)
17. Stanton, D.T.; Jurs, P.C. Development and use of charged partial surface area structural descriptors in computer-assisted quantitative structure-property relationship studies. *Anal. Chem.* **1990**, *62*, 2323–2329. [\[CrossRef\]](#)
18. Stanton, D.T.; Egolf, L.M.; Jurs, P.C.; Hicks, M.G. Computer-assisted prediction of normal boiling points of pyrans and pyrroles. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 306–316. [\[CrossRef\]](#)
19. Zhang, R.; Wang, Z.; Flisak, Z.; Hao, X.; Liu, Q.; Sun, W.-H. Achieving branched polyethylene waxes by aryliminocycloocta [b] pyridylnickel precatalysts: Synthesis, characterization, and ethylene polymerization. *J. Poly. Sci. Part. A Poly. Chem.* **2017**, *55*, 2601–2610. [\[CrossRef\]](#)
20. Sun, Z.; Huang, F.; Qu, M.; Yue, E.; Oleynik, I.V.; Oleynik, I.I.; Zeng, Y.; Liang, T.; Li, K.; Zhang, W.; et al. Targeting polyethylene waxes: 9-(2-cycloalkylphenylimino)-5, 6, 7, 8-tetrahydrocycloheptapyridylnickel halides and their use as catalysts for ethylene polymerization. *RSC Adv.* **2015**, *5*, 77913–77921. [\[CrossRef\]](#)
21. Yu, J.; Hu, X.; Zeng, Y.; Zhang, L.; Ni, C.; Hao, X.; Sun, W.-H. Synthesis, characterisation and ethylene oligomerization behaviour of N-(2-substituted-5, 6, 7-trihydroquinolin-8-ylidene) arylaminonickel dichlorides. *New J. Chem.* **2011**, *35*, 178–183. [\[CrossRef\]](#)
22. Sun, Z.; Yue, E.; Qu, M.; Oleynik, I.V.; Oleynik, I.I.; Li, K.; Liang, T.; Zhang, W.; Sun, W.-H. 8-(2-Cycloalkylphenylimino)-5, 6, 7-trihydro-quinolynickel halides: Polymerizing ethylene to highly branched and lower molecular weight polyethylenes. *Inorg. Chem. Front.* **2015**, *2*, 223–227. [\[CrossRef\]](#)
23. Wang, S.; Zhang, W.; Du, S.; Asuha, S.; Flisak, Z.; Sun, W.-H. Propyl substituted 4-arylimino-1, 2, 3-trihydroacridylnickel complexes: Their synthesis, characterization and catalytic behavior toward ethylene. *J. Organomet. Chem.* **2015**, *789*, 408–413. [\[CrossRef\]](#)
24. Wang, Z.; Zhang, Y.; Ma, Y.; Hu, X.; Solan, G.A.; Sun, Y.; Sun, W.-H. Molecular weight control of polyethylene waxes using a constrained imino-cyclopenta [b] pyridyl-nickel catalyst. *J. Poly. Sci. Part. A Poly. Chem.* **2017**, *55*, 3494–3505. [\[CrossRef\]](#)
25. Huang, F.; Sun, Z.; Du, S.; Yue, E.; Ba, J.; Hu, X.; Liang, T.; Galland, G.B.; Sun, W.-H. Ring-tension adjusted ethylene polymerization by aryliminocycloheptapyridyl nickel complexes. *Dalton Trans.* **2015**, *44*, 14281–14292. [\[CrossRef\]](#)
26. Winer, B.J.; Brown, D.R.; Michels, K.M. Single-Factor Experiments Having Repeated Measures on the Same Elements [Ch. 4]. In *Statistical Principles in Experimental Design*, 3rd ed.; McGraw-Hill: New York, NY, USA, 1991; pp. 220–281.
27. Jolliffe, I.T. Principal Component Analysis Series. In *Springer Series in Statistics*, 2nd ed.; Springer: New York, NY, USA, 2002; p. 487.
28. Delley, B. Ground-state enthalpies: Evaluation of electronic structure approaches with emphasis on the density functional method. *J. Phys. Chem. A* **2006**, *110*, 13632–13639. [\[CrossRef\]](#)
29. Delley, B. An all-electron numerical method for solving the local density functional for polyatomic molecules. *Chem. Phys.* **1990**, *92*, 508–517. [\[CrossRef\]](#)
30. Delley, B.J. From molecules to solids with the DMol3 approach. *Chem. Phys.* **2000**, *113*, 7756–7764. [\[CrossRef\]](#)
31. Perdew, J.P.; Wang, Y. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B.* **1992**, *45*, 13244–13249. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Becke, A.D. A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys.* **1988**, *88*, 2547–2553. [\[CrossRef\]](#)
33. Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. Energy-adjusted ab initio pseudopotentials for the first row transition elements. *J. Chem. Phys.* **1987**, *86*, 866–872. [\[CrossRef\]](#)
34. Bergner, A.; Dolg, M.; Küchle, W.; Stoll, H.; Preuß, H. Ab initio energy-adjusted pseudopotentials for elements of groups 13–17. *Mol. Phys.* **1993**, *80*, 1431–1441. [\[CrossRef\]](#)

-
35. Katritzky, A.R.; Lobanov, V.S.; Karelson, M. *Comprehensive Descriptors for Structural and Statistical Analysis (Codessa)*, Reference Manual; Semichem, Inc.: Shawnee Mission, KS, USA; Florida University: Gainesville, FL, USA, 2004.
 36. Frisch, M.J.E.A.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.; et al. *Gaussian 09*; Revision B.01; Gaussian, Inc.: Wallingford, CT, USA, 2010.
 37. Hehre, W.J.; Ditchfield, R.; Pople, J.A. Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261. [[CrossRef](#)]
 38. Gill, P.M.; Johnson, B.G.; Pople, J.A.; Frisch, M.J. An investigation of the performance of a hybrid of Hartree-Fock and density functional theory. *Int. J. Quantum Chem.* **1992**, *26*, 319–331. [[CrossRef](#)]
 39. Yi, J.; Yang, W.; Sun, W.-H. Quantitative Investigation of the Electronic and Steric Influences on Ethylene Oligo/Polymerization by 2-Azacycyl-6-aryliminopyridylmetal (Fe, Co, and Cr) Complexes. *Macromol. Chem. Phys.* **2016**, *217*, 757–764. [[CrossRef](#)]
 40. Yang, W.; Ma, Z.; Sun, W.-H. Modeling study on the catalytic activities of 2-imino-1, 10-phenanthrolineylmetal (Fe, Co, and Ni) precatalysts in ethylene oligomerization. *RSC Adv.* **2016**, *6*, 79335–79342. [[CrossRef](#)]
 41. Ahmed, S.; Yang, W.; Ma, Z.; Sun, W.-H. Catalytic activities of bis (pentamethylene) pyridyl (Fe/Co) complex analogues in ethylene polymerization by modeling method. *J. Phys. Chem. A* **2018**, *122*, 9637–9644. [[CrossRef](#)]
 42. Malik, A.A.; Yang, W.; Ma, Z.; Sun, W.-H. The Catalytic Activities of Carbocyclic Fused Pyridineimine Nickel Complexes Analogues in Ethylene Polymerization by Modeling Study. *Catalysts* **2019**, *9*, 520. [[CrossRef](#)]
 43. Roy, S.N. On a heuristic method of test construction and its use in multivariate analysis. *Ann. Math. Stat.* **1953**, *24*, 220–239. [[CrossRef](#)]
 44. Shao, J.J. Linear model selection by cross-validation. *J. Am. Stat. Assoc.* **1993**, *88*, 486–494. [[CrossRef](#)]
 45. Golbraikh, A.; Tropsha, A. Beware of q²! *J. Mol. Graph. Model.* **2002**, *20*, 269–276. [[CrossRef](#)]
 46. Chong, I.G.; Jun, C.H. Performance of some variable selection methods when multicollinearity is present. *Chemom. Intell. Lab. Syst.* **2005**, *78*, 103–112. [[CrossRef](#)]
 47. Gosselin, R.; Rodrigue, D.; Duchesne, C. A Bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications. *Chemom. Intell. Lab. Syst.* **2010**, *100*, 12–21. [[CrossRef](#)]
 48. Li, H.-D.; Xu, Q.-S.; Liang, Y.-Z. libPLS: An Integrated Library for Partial Least Squares Regression and Discriminant Analysis. *Chemom. Intell. Lab. Systems* **2018**, *176*, 34–43. Available online: <http://www.libpls.net> (accessed on 15 March 2020). [[CrossRef](#)]
 49. Martin, Y.C.; Kim, K.H.; Liu, C.T. Comparative Molecular Field Analysis: CoMFA. In *Advances in Quant. Str.-Property Relationships*; JAI Press: Greenwich, CT, USA, 1996; Volume 1, pp. 1–52.
 50. Cramer, R.D.; Bunce, J.D.; Patterson, D.E.; Frank, I.E. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967. [[CrossRef](#)] [[PubMed](#)]