The Journal of Engineering and Exact Sciences – JCEC, Vol. 05 N. 03 (2019)
journal homepage: https://periodicos.ufv.br/ojs/jcec
doi: 10.18540/jcecvl5iss3pp0283-0290
OPEN ACCESS – ISSN: 2527-1075

QSAR STUDY OF 2-SUBSTITUTED PHENYL-2-OXO-, 2-HYDROXYL- AND 2-ACYLLOXYETHYLSULFONAMIDES AS FUNGICIDES

Y. ISYAKU[1,*], A. UZAIRU[2] and S. UBA[3]
[1] Ahmadu Bello University, Department of Chemistry, Zaria, Kaduna State, Nigeria

[*]*Corresponding author Ahmadu Bello University, Department of Chemistry, Zaria, Kaduna State, Nigeria, Phone: +2348034984598*
e-mail addressl: isyakuyusuf01@gmail.com (Y. Isyaku).

A R T I C L E I N F O

Article history:
Received 2019-04-28
Accepted 2019-06-28
Available online 2019-06-30

k e y w o r d s
QSAR
2-substituted
Sulfonamides
Fungicides

A B S T R A C T

An insilico study was carried out on a series of thirty-five (35) sulfonyl-containing compounds for their antifungal activities against Botrytis Cinerea fungi using QSAR techniques. Using Spartan 14 molecular modelling software to draw the molecular structure of the compounds, the DFT/B3LYP/6-31G quantum method of the software was used in optimizing the drawn compounds. The optimized compounds of the dataset were then underbring into PaDEL-Descriptor software for their molecular descriptors calculation. The calculated PaDel-descriptors were then subjected to data-Pretreatment and later splitted into 70% training set and 30% test set. The model was generated using the training set and the test set for the validation of the model built. Using Genetic Function Algorithm (GFA) the model was developed. Four models were developed in which model 1 was chosen as the optimum model with good statistical parameters;* $R^2 = 0.954$, $R^2_{adj} = 0.941$, *cross validation* $R^2 / Q^2_{cv} = 0.888$ *and* $R^2_{pred} = 0.839$. *The model proposed was found to be stable, robust and showed a good internal and external validation. Other statistical analysis such as mean effect, variance inflation factor (VIF), Williams plot among others were also carried out for the applicability domain of the model.*

1. INTRODUCTION

Sulfonyl-containing compounds are one of the major active compounds categorized to have a vast range of biological activity. They are extensively used in drugs and agrochemicals. Botrytis Cinerea (scientifically) commonly known as Gray mold, Flower capsule blight, Botrytis brown stain, Scape blight or Bunch not, is a plant pathogen that infects over 200 plant species, causing grey mould which lead to serious economic losses of $10billion to $100billion annually (Boddy, Lynne, 2016). The plant species found to be affected by B. Cinerea includes tomatoes, Lettuce, Grapes, Strawberries e.t.c causing a grey powdery mould on the effected plants. It has the ability to counteract a large range of plant defence chemicals. It is one of the most extensively studied necrotrophict plant pathogens. B. Cinerea has some other relatives like B. byssoides, B. allii and B. squamosal which infect onions, B. Tulipae which infect Saffron and tulips, B. Fabae which affect beans and B. gladioli which infect gladioli and lilies. Due to development of several strains on many commercial antifugal compounds, the need of developing new antifungal with novel mode of action arised in order to specifically fight the resistance of these organism instead of the general fungicide. And also, to ensure the activity of the fungicides donot affect the beneficial organisms in the environment.

Sulfonyl-containing compounds excises a vital role in the field of agrochemicals as well as medicine. The first drugs that found to have selectivity on bacterial activity that could systematically be used to inhibit a specific bacterial infection was sulfonamides. Due to this great success, a considerably greater attention has been paid to develop more sulfonyl-containing compounds as agrochemicals and drugs. Some sulfonamides fungicides such as cyazofamid, tolnifamide and amisulbrom are commercially used. Sulfonyl-containing compounds such as sulfonamides drugs are used as anti-tumour (Huang *et al.,* 2001). Also, sulfonamides were found to have anti-plasmodial activity (Fisher *et al.,* 2017).

Sulfonyl-containing compounds such as sulfonamides are extensively used in pharmaceutical industries as anticancer, anti-inflammatory and antiviral agents. There are over 30 drugs containing this functionality of sulfonamides that are clinically used. This includes, antibacterial, anticonvulsant, diuretics, hypoglycemic and HIV protease inhibitors.

Due to pathogenic activity of Botrytis Cinerea organisms on both plants and animals and an extensively wide range of antibacterial activity by sulfonyl-containing compounds, many researches are carried out to fight the existence and pathogenic activity of the fungi. Some of these researches are computational studies such as QSAR study.

QSAR is a mathematical model which link the structure-derived characters of given compounds to their inhibitory activities. The studies of QSAR are intended at formulating a model (correlation models) using the activity and other informations from the chemical in the data in a statistical approch (Roy *et al.,* 2015). QSAR studies were carried out to predict more active compounds that will inhibit the activity of fungal diseases (Saiz-Urra *et al.,* 2009); (Singla *et al.,* 2009).

The aims of the paper are to develop a model (QSAR) which can predict a better activities of sulfonyl-containing compounds against Boytrytis Cinerea fungi.

2. MATERIALS AND METHOD

2.1 Data

Thirty-five (35) derivatives of 2-substituted phenyl-2-oxo-, 2-Hydroxy- and 2-Acyloxyethylsulfonamides used in the research are found in literature (Wang *et al.,* 2017). The activities of these compounds were reported in EC_{50} (mg/L), which were converted to pEC_{50} (pEC_{50} = -log1/ EC_{50}). The activity values and their corresponding molecular structure found in the date set are presented in the table 1 below.

3. TABLES 1- Compounds and pEC50 values

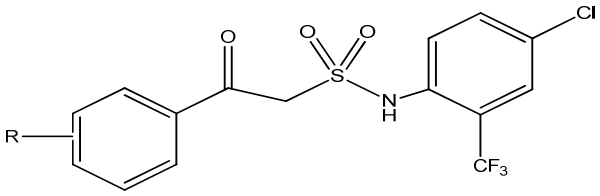

Figure 1- N-(2-trifluoromethyl-4-chlorophenyl)-2-substituted-phenyl-2-oxo-sulfonamides

Serial No.	R	pEC50
1	2-CH_3	1.210853
2	4-OCH_3	1.247973
3	2-F	0.909556
4	3-F	0.91169
5	2-Cl	0.899821
6	3-Cl	0.898176
7	4-Br	0.90309
8	2-CF_3	0.619093
9	3-CF_3	0.858537
10	4-CF_3	0.732394
11	4-NO_2	0.834421
12	3,4-F_2	0.838219
13	3,5-F_2	0.541579
14	2,4-Cl_2	0.598791
15	2-Cl-3-F	0.930949
16	3,5-(CF_3)	0.888741

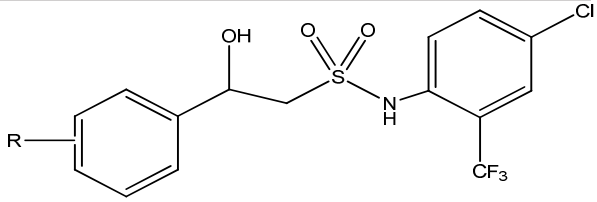

Figure 2- N-(2-trifluoromethyl-4-chlorophenyl)-2-substituted-2-hydroxy-sulfonamides

17	H	1.520615
18	3-F	0.855519
19	4-NO_2	0.609594
20	3,4-F_2	0.396199
21	3,5-F_2	0.21467

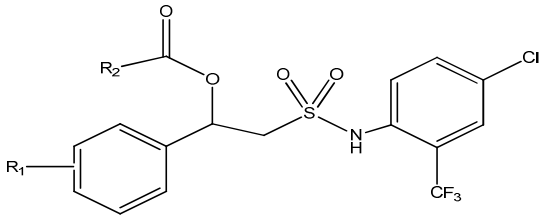

Figure3-N-(2-trifluoromethyl-4-chlorophenyl)-2-substituted-2-hydroxy-2-(3,5-difluorophenyl) ethylsulfonamides.

Serial No.	R^1	R^2	pEC50
22	3,5-F_2	CH_3	0.521138
23	3,5-F_2	2-$CH_3C_6H_4$	2.122281
24	3,5-F_2	3-$CH_3C_6H_4$	3.562114
25	3,5-F_2	4-$CH_3C_6H_4$	3.159281
26	3,5-F_2	4-$OCH_3C_6H_4$	0.991226
27	3,5-F_2	$ClCH_2CH_2$	1.139879
28	3,5-F_2	Cl_2CH	0.887617
29	3,5-F_2	Cl_3C	0.674861
30	3,5-F_2	2-FC_6H_4	2.076822
31	3,5-F_2	3-FC_6H_4	1.324077
32	3,5-F_2	2-ClC_6H_4	2.916717
33	3,5-F_2	3-ClC_6H_4	2.529341
34	3,5-F_2	2-$CF_3C_6H_4$	2.726915
35	3,5-F_2	3- $CF_3C_6H_4$	2.14473

Molecular Structure Optimization and Descriptors Calculation

Spartan 14 software (wavefunction,Inc. 2013) was used to optimized the drawn structures of the compounds where the Density Function Theory (DFT) version B3LYP corresponding to 6-31G* basis set was employed. In this process, all the molecular structures were drawn in the graphical user interface of the Spartan 14 software with the help of 2D application tool. These were later exported in the format "3D". The energy of the structures was minimized and then optimized i.e. calculate the quantum chemical descriptors, (Abdulfatai *et al.,* 2016).

Calculations of Molecular Descriptors

These are the properties of the molecule in numerical/mathematical values. PaDEL descriptor software version 2.18 was employed to calculate 1D, 2D and 3D descriptors.

Dataset Splitting

Dataset was splitted into training set and test set in the ratio of 3:1 by Kennard Stone Algorithm. The training set are the set of molecules that partakes in model building whereas test set are the unused data set and they are used to externally validate the model. The division is at 70% training set and 30% test set (Gramatica *et al.,* 2012).

QSAR Modelling and Validation

Using the training set the QSAR models was developed by employment of Genetic Function Approximation (GFA) available in the software (Material Studio, 2017). The models generated were internally validated while external validation of the built models through was carriedout by the employment of OECD principle 2007 using the test set. The nature of chemicals used in the training set for model generation influenced the predictive capacity of the built models, while the test set compounds would be well predicted as its molecules are similar to that of training set structurally (Roy and Mandal, 2008).

Parameters for Internal Validation of Model

(1) Friedman's Lack of Fit (LOF): this describe the fitness of the built model.

$$LOF = \frac{SEE}{\left(1-\frac{C+dP}{M}\right)^2} \quad \text{(i)}$$

where SEE = standard error estimation, C = number of terms present in the model, d = smooth parameter that is user-defined, P=total number of model's descriptors and M = number of training set molecules.

The regression model is in the for $Y = D_1x_1 + D_2x_2 + D_3x_3 \ldots . + D_nx_n + c$ just as the equation of straight line graph, $(Y = mx + c)$. Where Y represents the predicted activity, D is the corresponding coefficients, x is the independent variables and c is the regression constant (Ibrahim *et al.,* 2018).

(2) Total variation of the model R^2;

$$R^2 = 1 - \frac{\sum(Yexp - Yperd)^2}{\sum(Yexp - \bar{Y}train)^2} \quad \text{(ii)}$$

where Y_{exp} = observed/experimental activity, Y_{pred} = predicted activity/toxicity and $\bar{Y}_{train}$ = average observed activity for the training set (Adeniji *et al.,* 2018).

(3) For a suitable and trustworthy model R^2 should have an adjustment, therefore, R^2_{adj} is given by;

$$R^2_{adj} = (1-R^2)\frac{(n-1)}{n-P-1} = \frac{(n-1)(R^2-P)}{n-P+1} \quad \text{(iv)}$$

(Abdulfatai *et al.,* 2018).

(4) Cross-Validation coefficient (Q^2_{cv}), which describe the predictive power of the built model toward the activity of the new active compounds.

$$Q^2_{cv} = 1 - \frac{\sum(Ypred - Yexp)^2}{\sum(Yexp - \bar{Y}train)^2} \quad \text{(v)}$$

(Adedirin *et al.,* 2018).

External Validation

For model to be externally validated, we calculate the predicted R^2 as;

$$R^2 = 1 - \frac{\sum(Ypred - Yexp)^2}{\sum(Yexp - \bar{Y}train)^2} \quad \text{(vi)}$$

where $Y_{pred\ (test\ set)}$ and $Y_{exp\ (test\ set)}$ represents the values of predicted as well as the experimental activity of the test set and

$\bar{Y}_{train}$ is the average activities of the compounds in training set. (Abdullahi *et al.*, 2018).

Variance Inflation Factor (VIF)

To show that the model is well established, variance inflation factor is as well calculated which is defined by the equation $1/1\text{-}R^2$ where R^2 is the multiple correlation coefficient existing between the model's variables. If the VIF is equal to 1, it means that inter-correlation in each variable doesn't count, and if it ranges from 1 to 5, then is said to be suitable and acceptable. But if the VIF turn out to be greater than 10, this indicates the instability of the model and need to be reexamined. (Edache *et al.*, 2017); (Pourbasheer *et al.*, 2015).

Mean Effect (MF)

Mean effect is defined by the following;

$$Mean\ effect = \frac{B_j \sum_i^n D_j}{\sum_j^m (B_j \sum_i^n D_j)} \quad \text{(vii)}$$

where B_j and D_j are the j-descriptor coefficient in the model and the values of each descriptor in training set, while m and n stands for the number of molecular descriptors as well as number of molecules in a training set. To evaluate the significance of the model, the mean effect of each descriptor was calculated (Edache *et al.*, 2015).

Applicability Domain

Willian's plot was employed to examine the outliers and of course the swayful (influential) compounds and also to assert positively the robustness and confidence of the generated model. The William's plot was plotted using standardized residuals against the Leverage. In order to evaluate the model's applicability domain, the approach of leverage was employed. For a given chemical compounds, leverage is giving by the following equation;

$$h_i = X_i(X^T X)^{-1} X_i^T \quad \text{(viii)}$$

where h_i is the leverage of each compound, X_i is the training set compounds of the matrix i. X is the matrix of nxk descriptor in the training set molecules. X^T is the transpose of X-matrix.

The warning leverage (h*) defined as a boundary of normal values of an outlier X and is given by;

$$h^* = 3\frac{(d+1)}{m} \quad \text{(ix)}$$

The variable m stands for number of molecules in the training set and d is the descriptors describing the model.

2. RESULTS AND DISCUSSION

Descriptors calculation

QSAR was carried out to formulate a model which relate the structure of thirty-five (35) of sulfonyl-containing compounds (2-substituted phenyl-2-oxo-, 2-hyderoxy- and 2-acyloxyethylsulfonamides) with the respective activities to inhibit B. *Cinerea* fungi.

After optimizing the compounds in the dataset using Spartan 14 software, 32 quantum chemical descriptors were generated. These 32 descriptors were then combined with 1875 other descriptors obtained from PaDEL descriptor software giving a sum of 1907 descriptors.

Data Division

By employing Kennard-Stone method, the data was divided into training set 70% and test set 30% using the software "Data Division GUI 1.2".

Model and its Validation

Five descriptors were used in generating the model through the employment of Genetic Function Approximation (GFA) available in Material Studio Software. The equation pEC_{50} below represent the best model with its statistically validation parameters.

pEC_{50} = Y =13.308368320*FMF -0.338596475 * RNCS -0.012982836*TPSA -2.054638915*WD.unity +0.112869857* Wgamma2.volume -0.440310924.

The validation parameters shown in the table (2) below i.e. the highly calculated R^2 values (0.954), R^2_{adj} value (0.941) and R^2_{cv} value (0.888) of the selected model indicates that the model possess the acceptability criteria.

Table2 Validation parameter

Validation parameters	Model	QSAR Standard
Friedman LOF	**0.20012400**	**-**
R-squared	**0.95394800**	**≥ 6**
Adjusted R-squared	**0.94115600**	**-**
Cross validated R-squared	**0.88762300**	**≥ 5**
Significant Regression	**Yes**	**-**
Significant-of-regression F-value	**74.57284200**	**-**
Critical SOR F-value (95%)	**2.79410900**	**-**
Replicate points	**0**	**-**
Computed experimental error	**0.00000000**	**-**
Lack-of-fit points	**18**	**-**
Min expt. error for non-significant LOF (95%)	**0.16374300**	**-**

From the results of internal validation and that of the external validation [where the R-squares are 0.954 (internal) and 0.839 (external)] indicates a strong relationship between the observed and predicted activities. Additionally, the descriptors possesses of positive coefficient in the best chosen model '1' such as FMF (Complexity of a molecule) and Wgamma2.volume (Directional WHIM, weighted by Van der Waal's volumes) are to increase the inhibition activities of these compounds against B. *Cinerea* fungi while the negative once that is RNCS (Relative Negative Charge Surface area), TPSA (Sum of solvent accessible surface areas of atoms with absolute value of partial charges greater than or equal to 0.2) and WD.unity (Non-directional WHIM, weighted by unit weights) indicates that the inhibition activities of these compounds against B. *Cinerea* will be more when such descriptors reduces. Table (3) below is table of descriptions as well as the classes of descriptors that made up the built model.

S/N	Name	Description	Class
1	**FMF**	Complexity of a molecule	**2D**
2	**RNCS**	Relative Negative Charge Surface area	**3D**
3	**TPSA**	Sum of solvent accessible surface areas of atoms with absolute value of partial charges greater than or equal to 0.2	**3D**
4	**WD.unity**	Non-directional WHIM, weighted by unit weights	**3D**
5	**Wgamma2.vol.**	Directional WHIM, weighted by Van der Waal's volumes	**3D**

Table (4) present the external validation while table 5 is for the calculation of predicted R^2 of the model1.

Table 4

S/N	pEC50	FMF	RNCS	TPSA	WD.unity	Wgamma2.volume	Y_{pred}	$Y_{pred} - Y_{obs}$
8	0.619093	0.421053	3.121658	100.3648	0.825194	3.259162	1.475596	0.856503
10	0.732394	0.421053	4.499184	86.60491	0.829515	0.986989	0.922474	0.19008
11	0.834421	0.432432	5.201466	87.11753	0.822422	0.930122	0.837632	0.003211
12	0.838219	0.457143	4.846884	183.8772	0.77786	2.667684	0.318008	-0.52021
13	0.541579	0.457143	4.931783	185.1581	0.760518	2.74563	0.317062	-0.22452
14	0.598791	0.457143	5.825755	88.88254	0.818084	2.701184	1.141001	0.542211
19	0.609594	0.410256	3.961571	84.57743	0.74902	0.833283	1.135191	0.525597
20	0.396199	0.432432	3.978342	176.1183	0.756379	0.811185	0.218564	-0.17764
21	0.21467	0.432432	3.857209	180.447	0.720445	0.805483	0.276569	0.061899
22	0.521138	0.380952	0.26502	220.4838	0.776334	2.930685	0.413003	-0.10813
29	0.674861	0.380952	0.689922	212.84	0.693369	3.307903	0.581411	-0.09345

Table 5

S/N	$(Y_{pred} - Y_{obs})^2$	$Y_{MeanTrain}$	$Y_{pred} - Y_{MeanTrain}$	$(Y_{pred} - Y_{MeanTrain})^2$
8	0.733597	1.5674	-0.94831	0.899286
10	0.03613	1.5674	-0.83501	0.697235
11	1.03E-05	1.5674	-0.73298	0.537259
12	0.27062	1.5674	-0.72918	0.531705
13	0.050408	1.5674	-1.02582	1.052308
14	0.293993	1.5674	-0.96861	0.938204
19	0.276252	1.5674	-0.95781	0.917392
20	0.031554	1.5674	-1.1712	1.371711
21	0.003831	1.5674	-1.35273	1.829878
22	0.011693	1.5674	-1.04626	1.094664
29	0.008733	1.5674	-0.89254	0.796626
	$\sum(Y_{pred} - Y_{obs})^2 = 1.7168$			$\sum(Y_{pred} - Y_{MeanTrain})^2 = 10.6663$

$R^2 = (1-1.7168/10.6663) = 0.839044$

The observed and predicted activity of B.*Cinerea* inhibitors as a potential antifungal and their actual residual values are given in the table 6. This residual value is the different between the observed and predicted activities. The lower the residual values between the experimental and predicted activities signifies the higher prediction ability of the model.

Table 6: Comparison of experimental and predicted activity

Serial No.	pEC50	Predicted pEC50	Residual
1.	1.21085300	1.23667300	-0.02582000
2.	1.24797300	1.21079800	0.03717500
3.	0.90955600	0.98448500	-0.07492900
4.	0.91169000	0.71653900	0.19515100
5.	0.89982100	0.92473000	-0.02491000
6.	0.89817600	0.77638900	0.12178700
7.	0.90309000	1.16791400	-0.26482400
9.	0.85853700	0.84737200	0.01116500
15.	0.93094900	0.83007600	0.10087400
16.	0.88874100	1.15365400	-0.26491300
17.	1.52061500	1.41670600	0.10390900

18.	0.85551900	0.70043900	0.15508000
23.	2.12228100	2.08550200	0.03677900
24.	3.56211400	3.32908000	0.23303500
25.	3.15928100	3.43625600	-0.27697500
26.	0.99122600	1.17510800	-0.18388200
27.	1.13987900	0.81459200	0.32528700
28.	0.88761700	1.16564800	-0.27803000
30.	2.07682200	2.21789100	-0.14106900
31.	1.32407700	1.55761600	-0.23353900
32.	2.91671700	2.71851800	0.19820000
33.	2.52934100	2.32935600	0.19998400
34.	2.72691500	2.59043800	0.13647700
35.	2.14473000	2.23074100	-0.08601100

Descriptors Correlation Matrix

The descriptors of the chosen model (model1) was selected and performed a correlation matrix on them as shown in table 7. The values indicated that some descriptors are inter-correlated while some are not for their correlation coefficients are greater than 0.5. The variance inflation factor (VIF) values are within the range of 1 to 5 which indicated that the descriptors and model are suitable and acceptable.

Descriptors	FMF	RNCS	TPSA	WD.unity	Wgamma2.volume	VIF
FMF	1.0000					1.2942
RNCS	-0.1103	1.0000				4.1877
TPSA	0.3348	-0.8345	1.0000			4.1522
WD.unity	-0.1538	0.5017	-0.4809	1.0000		1.6420
Wgamma2.volume	-0.0575	-0.5312	0.4036	-0.5480	1.0000	1.7059

Some statistical parameters of the descriptors appeared in the built model are presented in the table 8 shown below. The magnitude of t-stat values for all descriptors are higher than 2 which signifies that the chosen descriptors are good (Adeniji *et al.*, 2018). Also all the descriptors has p-values of less than 0.05 which signifies good relation between the descriptors and the inhibition concentration of the compounds.

Descriptors	Coefficients	Standard error	t-stat	p-value	Mean effect
FMF	13.30837	1.409311	9.443172	2.14E-08	3.7879
RNCS	-0.3386	0.042274	-8.0095	2.41E-07	-0.5524
TPSA	-0.01298	0.002031	-6.39088	5.11E-06	-1.2873
WD.unity	-2.05464	0.289777	-7.09043	1.31E-06	-0.8852
Wgamma2.volume	0.11287	0.035541	3.175746	0.005234	0.2179

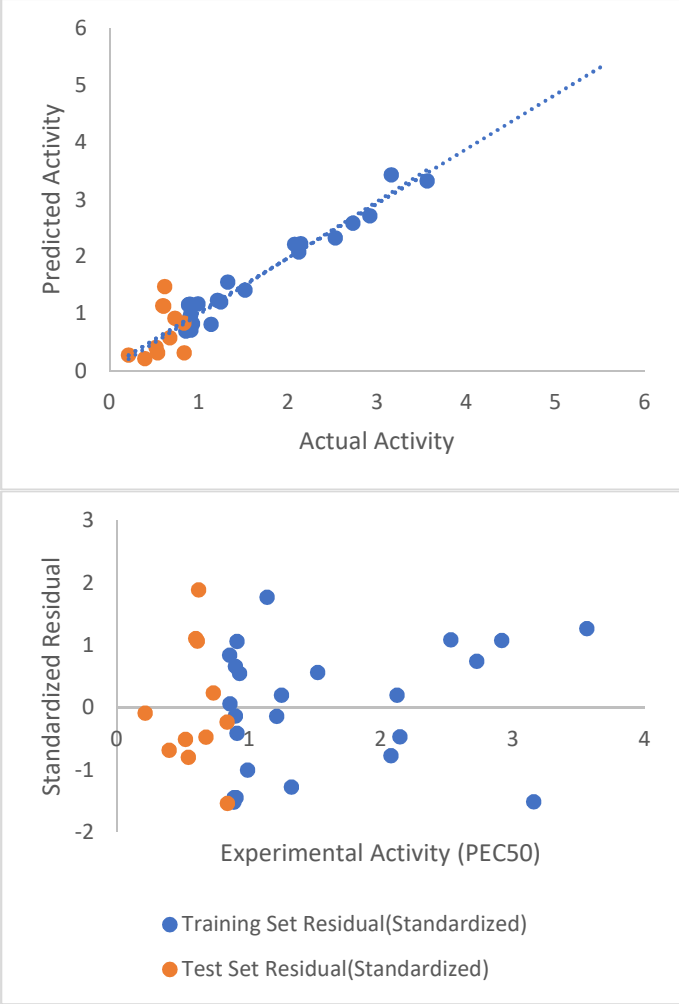

Figure 3 – Plot of standardized residual versus experimental activity (pEC_{50})

From figure 3, we witnessed a random disperse at a point where the standardized residual is zero which indicates the absence of systematic error while developing the model (Shola *et al.*, 2018).

Wiiliam's Plot of Model 1

A Williams plot as shown in figure 4, is a plot of standardized residual versus leverages (for both the training set and test set compounds) of built model. The essence of this plot is to examine the presence of an outliers together with other influencing molecules present in the model. The result revealed that two (2) compounds from the test set were outside applicability domain of the compounds which indicated that the two compounds may have different structure from other in the dataset. Hence the compounds are beyond the threshold value or warning leverages h* which was calculated to be 0.75.

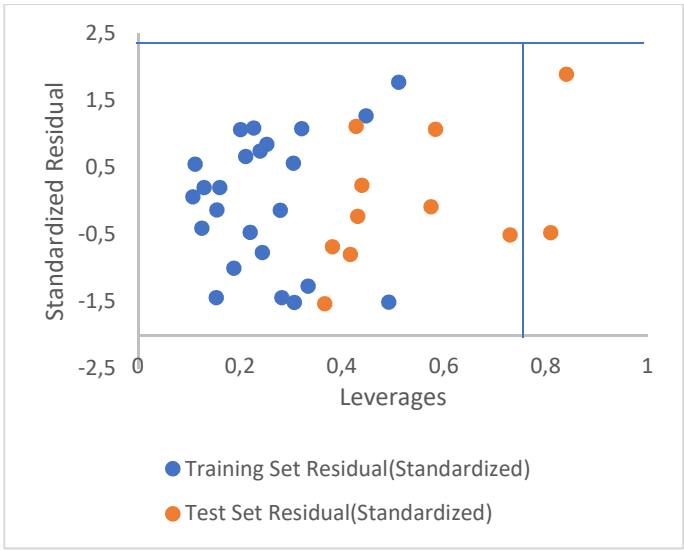

Figure 4 – Williams plot: a plot of standardized residual versus leverages (for both the training set and test set).

Conclusion

The QSAR model for 2-substituted phenyl-2-oxo-, 2-Hydroxy- and 2-Acyloxyethylsulfonamides was successfully developed which predicted the toxic activity of the compounds against B. *Cinerea* by employing Genetic Function Approximation method. With model 1 being the best model, the R^2, R^2_{adj} and Q^2_{cv} are 0.954, 0.941 and 0.888 respectively, and the external validation R^2_{pred} = 0.839. The research found that the toxicity of compounds was as a result of the molecular descriptors FMF, RNCS, TPSA, WD. unity and Wgamma2.volume with their mean effect values of 3.787873, 0.55237, -1.28729, -0.88524 and 0.217949 respectively. This finding provides a guideline for development of new/novel sulfonyl compounds with excellent toxicity against B. *Cinera* fungi. Some of these compounds may include compounds 4, 6 and 18 (with pEC_{50} of 0.71654, 0.77639 and 0.70044).

REFERENCES

Abdulfatai, Usman, Adamu Uzairu, and Sani Uba. "Quantitative structure activity relationship study of anticonvulsant activity of α_substituted acetamido-N-benzylacetamide derivatives." *Cogent Chemistry* 2, 2016.

Abdulfatai, Usman, Adamu Uzairu, and Sani Uba. "Investigation of some acetamido-N-benzylacetamide derivatives as potent anti-convulsant inhibitors." *Journal of Computational Methods in Molecular Design* 5, 2015.

Mustapha, Abdullahi, Gideon Shallangwa, Muhammad Tukur Ibrahim, Abdullahi Umar Bello, David Arthur Ebuka, Adamu Uzairu, and Paul Mamza. "QSAR studies on some C14-urea tetrandrine compounds as potent anti-cancer against Leukemia cell line (K562)." *Journal of the Turkish Chemical Society, Section A: Chemistry* 5, 2018.

Adeniji, Shola Elijah, David Ebuka Arthur, and Adedirin Oluwaseye. "Computational modeling of 4-

Phenoxynicotinamide and 4-Phenoxypyrimidine-5-carboxamide derivatives as potent anti-diabetic agent against TGR5 receptor." *Journal of King Saud University-Science*, 2018.

Adeniji, Shola Elijah, Sani Uba, and Adamu Uzairu. "QSAR Modeling and Molecular Docking Analysis of Some Active Compounds against Mycobacterium tuberculosis Receptor (Mtb CYP121)." *Journal of pathogens* (2018), 2018.

Adeniji, Shola Elijah, Sani Uba, and Adamu Uzairu. "In silico study for investigating and predicting the activities of 1, 2, 4-triazole derivaties as potent anti-tubercular agents." *The Journal of Engineering and Exact Sciences* 4, 2018.

Antifungal Activity Relationships of Clovane Derivatives against Botrytis cinerea." *Journal of agricultural and food chemistry* 57, 2009.

Arthur, David Ebuka. "Toxicity modelling of some active compounds against k562 cancer cell line using genetic algorithm-multiple linear regressions." *Journal of the Turkish Chemical Society, Section A: Chemistry* 4, 2016.

Boddy, Lynne. "Pathogens of Autotrophs." In *The Fungi*, pp. 245-292. Academic Press, 2016.

Dehmer, Matthias, Kurt Varmuza, Danail Bonchev, and Frank Emmert-Streib, eds. *Statistical modelling of molecular descriptors in QSAR/QSPR*. Weinheim, Germany: Wiley-VCH, 2012.

Edache, Emmanuel Israel, David Ebuka Arthur, and Usman Abdulfatai. "tyrosine Activity of Some Tetraketone and Benzyl-benzoate Derivatives Based on Genetic Algorithm-Multiple Linear Regression." 2017.

Fisher, Robert S., J. Helen Cross, Carol D'souza, Jacqueline A. French, Sheryl R. Haut, Norimichi Higurashi, Edouard Hirsch et al. "Instruction manual for the ILAE 2017 operational classification of seizure types." *Epilepsia* 58, 2017.

Gramatica, Paola, Stefano Cassani, Partha Pratim Roy, Simona kovarich, Chun Wie Yap, Ester Papa. "QSAR modeling is not "push button and find a correlation": a case study of toxicity of (benzo-)triazoles on algae" *Molecular Informatics* 31, 2012.

Huang, Zhaohua, Zhaoliang Lin, and Junlian Huang. "A novel kind of antitumour drugs using sulfonamide as parent compound." European journal of medicinal chemistry 36, 2001.

Ibrahim, Muhammad Tukur, Adamu Uzairu, Gideon Adamu Shallangwa, and Abdulkadir Ibrahim. "Computational studies of some biscoumarin and biscoumarin thiourea derivatives as α-glucosidase inhibitors." *The Journal of Engineering and Exact Sciences* 4, 2018.

Ibrahim, Muhammad Tukur, Adamu Uzairu, Gideon Adamu Shallangwa, and Abdulqadir Ibrahim. "In-silico studies of some oxadiazoles derivatives as anti-diabetic compounds." *Journal of King Saud University-Science*, 2018.

Pourbasheer, Eslam, Reza Aalizadeh, Mohammad Reza Ganjali, and Parviz Norouzi. "QSAR study of IKKβ inhibitors by the genetic algorithm: multiple linear regressions." *Medicinal Chemistry Research* 23, 2014.

Roy, Kunal, and Asim Sattwa Mandal. "Development of linear and nonlinear predictive QSAR models and their external validation using molecular similarity principle for anti-HIV indolyl aryl sulfones." *Journal of enzyme inhibition and medicinal chemistry* 23, 2008.

Roy, Kunal, Supratik Kar, and Rudra Narayan Das. A primer on QSAR/QSPR modeling: Fundamental concepts. Springer, 2015.

Singla, Rajeev K., and Varadaraj Bhat G. "QSAR model for predicting the fungucidal action of 1,2,4-triazole derivatives against candida albicans." Journal of enzyme inhibition and medicinal chemistry 25, 2010.

Singla, Rajeev K., and Varadaraj Bhat G. "QSAR model for predicting the fungicidal action of 1, 2, 4-triazole derivatives against Candida albicans." *Journal of enzyme inhibition and medicinal chemistry* 25, 2010.

Saiz-Urra, Liane, Juan C. Racero, Antonio J. Macias-Sanchez, Rosario Hernandez-Galan, James R. Hanson, Maykel Perez-Gonzalez, and Isidro G. Collado. "Synthesis and Quantitative Structure− Antifungal Activity Relationships of Clovane Derivatives against Botrytis cinerea." *Journal of agricultural and food chemistry* 57, 2009.

Wang, Minlong, Peng Rui, Caixiu Liu, Ying Du, Peiwen Qin, Zhiqiu Qi, Mingshan Ji, Xinghai Li, and Zining Cui. "Design, synthesis and fungicidal activity of 2-substituted phenyl-2-oxo-, 2-hydroxy-and 2-acyloxyethylsulfonamides." *Molecules* 22, 2017.

Wavefunction, Inc. "Spartan'14, version 1.1. 2." 2013.

Yap, Chun Wei. "PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints." *Journal of computational chemistry* 32, 2011.