

**Estudo de Modelos QSAR de uma Série de Inibidores da Desidroquinato Desidratase II de *Helicobacter pylori***

QSAR Models Study of a Series of Dehydroquinone Dehydratase II Inhibitors of *Helicobacter pylori*

**Paulo H. de S. Miranda,<sup>a</sup> Pedro I. C. de Oliveira,<sup>b</sup> Estela M. G. Lourenço,<sup>a,c</sup> Alexander M. S. Moraes,<sup>a</sup> Priscilla S. de S. N. Silverio,<sup>b</sup> Euzébio G. Barbosa<sup>a,b</sup> Alessandro K. Jordão<sup>a,\*</sup>**

*a Universidade Federal do Rio Grande do Norte, Programa de Pós-Graduação em Ciências Farmacêuticas, Natal-RN, Brasil*

*b Universidade Federal do Rio Grande do Norte, Programa de Pós-Graduação em Bioinformática, Natal-RN, Brasil*

*c Universidade Federal de Minas Gerais, Laboratório de Modelagem Molecular e Planejamento de Fármacos, Belo Horizonte-MG, Brasil*

\* [alessandro.jordao@ufrn.br](mailto:alessandro.jordao@ufrn.br)

## 1. Dados Estatísticos Adicionais Sobre O Modelo Híbrido

**Tabela S1.** Dados estatísticos adicionais sobre o modelo híbrido

| Variável | Erro Padrão<br>(std err) | Valor t (t) | P-valor<br>(P> t ) | Intervalo de<br>Confiança [0.025<br>- 0.975] | VIF    |
|----------|--------------------------|-------------|--------------------|----------------------------------------------|--------|
| HF+      | 0,006                    | 2,361       | 0,025              | 0,002 – 0,026                                | 3,523  |
| HF-      | 0,315                    | -2,821      | 0,009              | -1,534 – -0,243                              | 16,252 |
| HF-      | 0,209                    | -3,756      | 0,001              | -1,211 – -0,356                              | 1,45   |
| AATS8i   | 0,002                    | 4,171       | < 0,001            | 0,004 – 0,012                                | 18,709 |
| QQ       | 1,019                    | 3,824       | 0,001              | 1,810 – 5,986                                | 1,247  |
| nHBDon   | 0,236                    | -5,278      | < 0,001            | -1,730 – -0,763                              | 40,427 |
| nsssCH   | 0,193                    | -2,705      | 0,012              | -0,918 – -0,127                              | 46,682 |
| LJ       | 0,006                    | -2,566      | 0,016              | -0,027 – -0,003                              | 1,643  |

O modelo de regressão linear múltipla foi ajustado aos dados de 37 observações para investigar a relação entre oito variáveis preditoras e a variável dependente. O modelo geral mostrou-se estatisticamente

significativo ( $F = 22,16$ ,  $p < 0,001$ ), indicando que o conjunto de preditores explica uma porção significativa da variância da atividade biológica.

A análise de multicolinearidade, realizada através do cálculo do Fator de Inflação de Variância (VIF, Tabela 2), revelou a presença de correlações lineares substanciais entre alguns preditores. Notavelmente, as variáveis nHBDon (VIF  $\approx 40,4$ ), nsssCH (VIF  $\approx 46,7$ ), AATS8i (VIF  $\approx 18,7$ ), e HF\_31,96450\_28,94940\_39,78190 (VIF  $\approx 16,3$ ) exibiram valores de VIF marcadamente elevados, ultrapassando os limiares convencionalmente aceitos (e.g., 5 ou 10). Tecnicamente, isso implica que a variância dos coeficientes estimados para essas variáveis é inflada devido à sua intercorrelação. As implicações primárias são a instabilidade das estimativas pontuais dos coeficientes (Tabela 1) e o aumento dos seus erros padrão, dificultando a inferência precisa sobre o efeito marginal isolado de cada um desses preditores específicos sobre a atividade biológica. Contudo, é fundamental ressaltar que a presença de multicolinearidade não invalida, por si só, a utilidade preditiva do modelo. Dado o elevado coeficiente de determinação ajustado ( $R^2$  ajustado = 0,825) e a forte significância global do modelo (Teste F,  $p < 0,001$ ), o conjunto de preditores demonstra um poder preditivo conjunto robusto. Portanto, o modelo mantém seu valor para fins de previsão.

## **2. Modelo Híbrido**

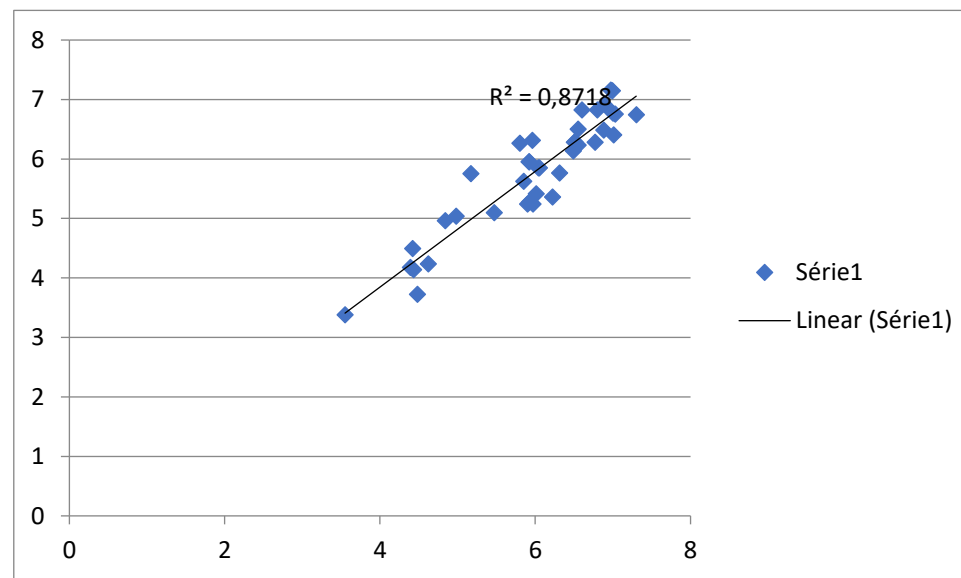
### **2.1. *Treinamento***

**Tabela S2.** Descritores moleculares utilizados no modelo híbrido

| Code  | HF_25.96450_23.94940_43.78190 | HF_31.96450_28.94940_39.78190 | AATS8i         | HF_22.96450_31.94940_37.78190 | QQ_30.96450_27.94940_40.78190 | nHBDon | nsssCH | LJ_23.96450_27.94940_40.78190 | Y.EXP | Y.PRED   |
|-------|-------------------------------|-------------------------------|----------------|-------------------------------|-------------------------------|--------|--------|-------------------------------|-------|----------|
| I F   | -0,0115                       | 0,899                         | 0<br>158,60    | -0,22799                      | -0,03311                      | 4      | 2      | -0,07379                      | 3,55  | 3,383136 |
| I C   | -0,02971                      | 0,94285                       | 838<br>183,49  | -0,19993                      | -0,1137                       | 3      | 3      | 31,84821                      | 4,39  | 4,178774 |
| I S   | -0,08039                      | 0,97572                       | 8823<br>159,09 | 1,20448                       | -0,00835                      | 3      | 2      | 32,81677                      | 4,42  | 4,495099 |
| I E   | -0,01907                      | 0,68945                       | 5588<br>159,09 | -0,22176                      | -0,02673                      | 4      | 3      | 1,71198                       | 4,43  | 4,143213 |
| I J   | -0,01094                      | 0,97987                       | 5588<br>158,06 | -0,24625                      | -0,02256                      | 4      | 3      | 6,32001                       | 4,48  | 3,725713 |
| I I   | 2,32029                       | 0,97777                       | 5022<br>158,60 | 1,41166                       | -0,13215                      | 3      | 2      | -0,15933                      | 4,62  | 4,240485 |
| I D   | -0,01089                      | 0,8148                        | 838<br>149,07  | 0,43851                       | 0,02285                       | 3      | 3      | 0,99516                       | 4,98  | 5,042472 |
| IV E  | 32,39884                      | 0,98122                       | 7474<br>183,76 | -0,21452                      | -0,15711                      | 3      | 2      | 5,58135                       | 5,17  | 5,756774 |
| III G | -0,05901                      | 0,72861                       | 4888<br>184,91 | -0,25664                      | -0,0111                       | 3      | 3      | 31,49231                      | 5,47  | 5,099188 |
| III C | -0,01845                      | 0,61519                       | 7652<br>167,96 | -0,22403                      | -0,002                        | 3      | 2      | 0,6601                        | 5,8   | 6,268541 |
| II E  | 1,49841                       | 0,6635                        | 3363<br>169,61 | -0,24661                      | -0,00265                      | 3      | 3      | -0,13369                      | 5,85  | 5,628469 |
| I T   | -0,02695                      | 0,88255                       | 4496<br>157,43 | -0,1984                       | -0,01627                      | 3      | 2      | 34,50659                      | 5,9   | 5,24188  |
| I A   | 34,07029                      | 0,86127                       | 3224<br>168,86 | -0,27156                      | 0,08147                       | 4      | 2      | -0,06532                      | 5,92  | 5,955    |
| II H  | -0,077                        | 0,75381                       | 5205<br>152,78 | -0,05079                      | -0,02133                      | 3      | 3      | -0,14517                      | 5,93  | 5,28185  |
| IV A  | 32,36737                      | 0,6969                        | 5054<br>158,26 | -0,22135                      | -0,0302                       | 3      | 2      | 31,58377                      | 5,96  | 6,317814 |
| III E | -0,00566                      | 0,73423                       | 9154           | -0,22425                      | -0,03087                      | 3      | 3      | 3,38696                       | 5,97  | 5,243445 |

|       |          |         |        |          |          |   |   |          |      |          |
|-------|----------|---------|--------|----------|----------|---|---|----------|------|----------|
| II D  | 35,72512 | 0,99977 | 158,60 | 0,01172  | 0,06976  | 3 | 2 | -0,0716  | 6,8  | 6,828208 |
|       |          |         | 838    |          |          |   |   |          |      |          |
| II I  | -0,12096 | 0,71417 | 165,23 | -0,24658 | -0,02133 | 3 | 3 | -0,15026 | 6,01 | 5,416953 |
|       |          |         | 1261   |          |          |   |   |          |      |          |
| IV O  | 32,11701 | 0,98197 | 156,67 | -0,19778 | -0,1619  | 3 | 2 | -0,1197  | 6,04 | 5,85102  |
|       |          |         | 2059   |          |          |   |   |          |      |          |
| III A | -0,01167 | 0,55532 | 162,80 | 0,01332  | -0,02462 | 3 | 3 | 6,19423  | 6,22 | 5,361719 |
|       |          |         | 9463   |          |          |   |   |          |      |          |
| I B   | 2,52144  | 0,91286 | 159,09 | -0,24882 | -0,01299 | 3 | 2 | -0,16908 | 6,31 | 5,765002 |
|       |          |         | 2993   |          |          |   |   |          |      |          |
| IV C  | -0,03204 | 0,93211 | 169,10 | 0,16977  | -0,01923 | 3 | 2 | 34,02571 | 4,84 | 4,965398 |
|       |          |         | 5154   |          |          |   |   |          |      |          |
| III I | -0,06611 | 0,6642  | 183,82 | -0,17427 | -0,01184 | 3 | 2 | -0,29818 | 6,49 | 6,140317 |
|       |          |         | 893    |          |          |   |   |          |      |          |
| I P   | 33,27201 | 0,99852 | 156,67 | -0,2028  | -0,06472 | 3 | 2 | -0,065   | 6,5  | 6,283934 |
|       |          |         | 2059   |          |          |   |   |          |      |          |
| IV M  | 32,1453  | 0,9973  | 154,81 | -0,20993 | -0,05522 | 3 | 2 | -0,11401 | 6,51 | 6,286368 |
|       |          |         | 5762   |          |          |   |   |          |      |          |
| IV D  | 32,39426 | 0,79718 | 154,06 | -0,13745 | -0,04321 | 3 | 2 | 3,82655  | 6,55 | 6,504729 |
|       |          |         | 5455   |          |          |   |   |          |      |          |
| IV L  | 31,9444  | 0,98814 | 153,48 | -0,21673 | -0,05995 | 3 | 2 | 2,14846  | 6,55 | 6,233741 |
|       |          |         | 0157   |          |          |   |   |          |      |          |
| IV G  | 31,28676 | 0,00051 | 155,01 | -0,28794 | -0,21656 | 3 | 2 | -0,08177 | 6,89 | 6,91656  |
|       |          |         | 3897   |          |          |   |   |          |      |          |
| I K   | 35,72512 | 0,99977 | 158,60 | 0,01172  | 0,06976  | 3 | 2 | -0,0716  | 6,6  | 6,828208 |
|       |          |         | 838    |          |          |   |   |          |      |          |
| IV J  | 31,85911 | 0,93916 | 147,93 | -0,22116 | -0,01506 | 3 | 2 | -0,11473 | 6,88 | 6,489409 |
|       |          |         | 24     |          |          |   |   |          |      |          |
| IV H  | -0,10326 | 0,46388 | 172,28 | -0,22653 | 0,00025  | 3 | 2 | -0,13254 | 7,01 | 6,406221 |
|       |          |         | 5429   |          |          |   |   |          |      |          |
| I L   | 35,16241 | 0,99994 | 160,23 | -0,2317  | 0,11775  | 3 | 2 | -0,06951 | 6,97 | 7,162732 |
|       |          |         | 3414   |          |          |   |   |          |      |          |
| I N   | 34,84509 | 0,99439 | 159,91 | -0,20507 | 0,11932  | 3 | 2 | -0,06887 | 6,99 | 7,150861 |
|       |          |         | 987    |          |          |   |   |          |      |          |

|      |          |         |                |          |          |   |   |          |      |          |
|------|----------|---------|----------------|----------|----------|---|---|----------|------|----------|
| IV F | 0        | 0       | 154,76<br>0659 | 0        | -0,00069 | 3 | 2 | 0        | 7    | 6,774459 |
| I Q  | 34,49794 | 0,80178 | 150,39<br>5703 | -0,21403 | -0,01544 | 3 | 2 | -0,06498 | 7,03 | 6,756029 |
| IV I | 32,20329 | 0,66331 | 152,52<br>4066 | -0,23029 | -0,04368 | 3 | 2 | 1,72731  | 7,3  | 6,74497  |
| II C | 33,27201 | 0,99852 | 156,67<br>2059 | -0,2028  | -0,06472 | 3 | 2 | -0,065   | 6,77 | 6,283934 |
| II K | 32,11701 | 0,98197 | 156,67<br>2059 | -0,19778 | -0,1619  | 3 | 2 | -0,1197  | 6,05 | 5,85102  |
| I R  | 34,07029 | 0,86127 | 157,43<br>3224 | -0,27156 | 0,08147  | 4 | 2 | -0,06532 | 5,92 | 5,955    |

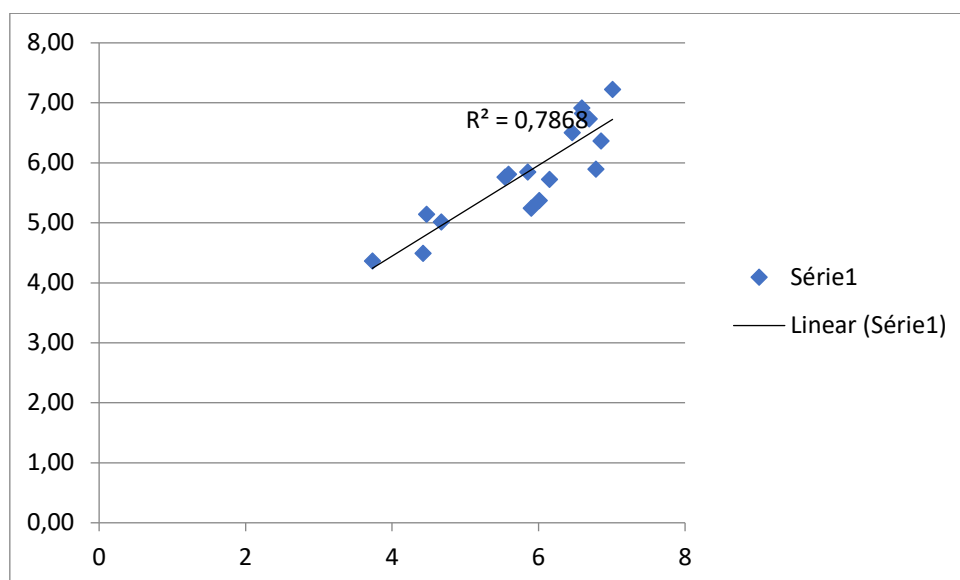


**Figura S1.** Previsão da atividade biológica com o modelo híbrido com coeficiente de determinação  $R^2$

## 2.2. Teste

**Tabela S3.** Descritores moleculares utilizados no modelo híbrido para o conjunto teste

| Code  | HF_25.96450_<br>23.94940_43.<br>78190 | HF_31.96450_<br>28.94940_39.7<br>8190 | AATS8i   | HF_22.96450<br>_31.94940_37<br>.78190 | QQ_30.96450_2<br>7.94940_40.781<br>90 | nHBDon | nsssCH | LJ_23.96450_2<br>7.94940_40.78<br>190 | Y.EXP | Y.PRED |
|-------|---------------------------------------|---------------------------------------|----------|---------------------------------------|---------------------------------------|--------|--------|---------------------------------------|-------|--------|
| I G   | -0,03678                              | 0,99899                               | 185,0065 | -0,19898                              | -0,046                                | 4      | 2      | -0,09669                              | 3,73  | 4,37   |
| III F | -0,03673                              | 0,82229                               | 157,5454 | 0,21385                               | -0,01685                              | 3      | 3      | 33,42405                              | 4,42  | 4,50   |
| III B | -0,01327                              | 0,99835                               | 181,0099 | -0,26039                              | -0,02066                              | 3      | 3      | 0,79106                               | 4,47  | 5,15   |
| II A  | -0,01004                              | 0,71353                               | 181,0099 | -0,25568                              | -0,13965                              | 3      | 3      | -0,29447                              | 4,67  | 5,02   |
| I H   | 5,57867                               | 0,98084                               | 157,6977 | -0,26231                              | -0,01207                              | 3      | 2      | -0,10642                              | 5,54  | 5,77   |
| III D | -0,02759                              | 0,54314                               | 168,8652 | -0,22861                              | -0,00365                              | 3      | 3      | 32,75563                              | 5,9   | 5,25   |
| II G  | 0,78329                               | 0,90733                               | 155,0014 | -0,27365                              | 0,01135                               | 3      | 2      | 0,04564                               | 5,59  | 5,81   |
| II B  | 0,02317                               | 0,48721                               | 200,4974 | -0,22249                              | -0,15126                              | 3      | 3      | -0,25003                              | 6,01  | 5,37   |
| I O   | 33,64144                              | 0,99487                               | 163,3618 | -0,23161                              | 0,06883                               | 4      | 2      | -0,06517                              | 6,15  | 5,73   |
| III H | 32,39426                              | 0,79718                               | 154,0655 | -0,13745                              | -0,04321                              | 3      | 2      | 3,82655                               | 6,46  | 6,50   |
| IV K  | 32,04242                              | 0,70064                               | 147,7903 | -0,24229                              | -0,03416                              | 3      | 2      | -0,11344                              | 6,69  | 6,74   |
| IV N  | 31,75529                              | 0,97678                               | 153,4169 | 0,462                                 | -0,06129                              | 3      | 2      | -0,11416                              | 6,78  | 5,90   |
| IV P  | 31,11712                              | 0,69474                               | 154,0919 | -0,00898                              | -0,08738                              | 3      | 2      | 2,16684                               | 6,85  | 6,37   |
| I M   | 35,29332                              | 0,99991                               | 159,2786 | -0,20212                              | 0,13677                               | 3      | 2      | -0,07166                              | 7,01  | 7,23   |
| IV B  | 31,28676                              | 0,00051                               | 155,0139 | -0,28794                              | -0,21656                              | 3      | 2      | -0,08177                              | 6,59  | 6,92   |
| II F  | 35,72512                              | 0,99977                               | 158,6084 | 0,01172                               | 0,06976                               | 3      | 2      | -0,0716                               | 6,6   | 6,83   |
| II J  | 32,11701                              | 0,98197                               | 156,6721 | -0,19778                              | -0,1619                               | 3      | 2      | -0,1197                               | 5,85  | 5,85   |



**Figura S2.** Previsão da atividade biológica com o modelo híbrido com coeficiente de determinação  $R^2$  para o conjunto teste

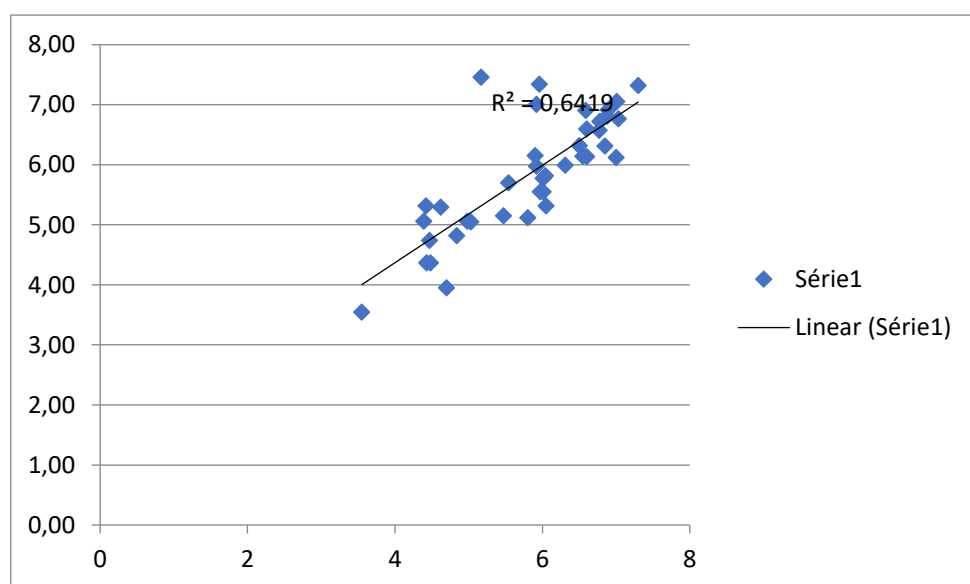
### 3. Modelo 2D

#### 3.1. Treinamento

**Tabela S4.** Valores dos descritores 2D do modelo 2D

| Code  | topoRadius | nHBint5 | nBondsM | GATS4c     | piPC5      | Y.EXP | Y.PRED |
|-------|------------|---------|---------|------------|------------|-------|--------|
| I A   | 6          | 3       | 14      | 1,1114488  | 5,65336599 | 5,92  | 5,98   |
| I K   | 5          | 3       | 9       | 1,14749899 | 5,32026231 | 6,6   | 6,14   |
| I H   | 6          | 3       | 15      | 1,09490998 | 5,75811282 | 5,54  | 5,70   |
| I M   | 6          | 3       | 10      | 1,10565503 | 5,51468803 | 7,01  | 7,05   |
| I J   | 5          | 7       | 8       | 1,04227822 | 5,02675468 | 4,48  | 4,37   |
| IV B  | 8          | 3       | 24      | 1,24203985 | 6,38982112 | 6,59  | 6,91   |
| I R   | 6          | 3       | 11      | 1,09332734 | 5,76773489 | 5,92  | 7,01   |
| III C | 4          | 3       | 2       | 1,10303812 | 3,98898405 | 5,8   | 5,12   |
| IV D  | 6          | 4       | 17      | 1,31679526 | 5,83663655 | 6,55  | 6,14   |
| III D | 5          | 3       | 7       | 1,08186719 | 5,13211519 | 5,9   | 6,16   |
| IV Q  | 5          | 3       | 7       | 1,16342656 | 4,54727621 | 6,04  | 5,82   |
| I B   | 6          | 3       | 14      | 1,11461548 | 5,65336599 | 6,31  | 6,00   |
| III B | 4          | 3       | 6       | 1,07099286 | 4,79579055 | 4,47  | 4,74   |
| IV I  | 8          | 4       | 22      | 1,25887293 | 6,3493575  | 7,3   | 7,32   |
| III G | 4          | 3       | 6       | 1,05878478 | 5,14531186 | 5,47  | 5,15   |
| IV J  | 6          | 3       | 12      | 1,15395117 | 5,61038713 | 6,88  | 6,81   |
| IV F  | 6          | 3       | 17      | 1,2704049  | 5,81170221 | 7     | 6,12   |
| IV G  | 6          | 3       | 12      | 1,15121688 | 5,69897929 | 6,89  | 6,92   |
| IV A  | 6          | 3       | 13      | 1,27899147 | 5,72113111 | 5,96  | 7,34   |
| I Q   | 6          | 3       | 13      | 1,14271805 | 5,86149822 | 7,03  | 6,76   |
| IV C  | 4          | 3       | 3       | 1,08192133 | 4,09434456 | 4,84  | 4,82   |
| IV E  | 8          | 4       | 22      | 1,28336085 | 6,3493575  | 5,17  | 7,46   |
| II B  | 5          | 3       | 8       | 1,09472552 | 5,03897977 | 6,01  | 5,78   |

|       |   |   |    |            |            |      |      |
|-------|---|---|----|------------|------------|------|------|
| II C  | 5 | 3 | 7  | 1,14000894 | 5,19642306 | 6,77 | 6,57 |
| I P   | 5 | 3 | 8  | 1,1838556  | 5,06811773 | 6,5  | 6,32 |
| II F  | 5 | 3 | 6  | 1,1091118  | 5,10518761 | 6,6  | 6,60 |
| I D   | 5 | 5 | 8  | 1,0678855  | 5,02675468 | 4,98 | 5,06 |
| I E   | 5 | 7 | 8  | 1,04227822 | 5,02675468 | 4,43 | 4,37 |
| IV P  | 7 | 3 | 22 | 1,2773497  | 6,22381558 | 6,85 | 6,31 |
| I T   | 4 | 3 | 7  | 1,08259971 | 5,20880286 | 5,03 | 5,05 |
| IV N  | 6 | 3 | 12 | 1,14646077 | 5,57902203 | 6,78 | 6,72 |
| II I  | 5 | 5 | 8  | 1,03225242 | 5,5346239  | 6,01 | 5,55 |
| I F   | 4 | 7 | 2  | 0,9930899  | 4,09434456 | 3,55 | 3,55 |
| I C   | 5 | 5 | 8  | 1,0678855  | 5,02675468 | 4,39 | 5,06 |
| III E | 5 | 5 | 8  | 1,03225242 | 5,5346239  | 5,97 | 5,55 |
| III F | 5 | 3 | 12 | 1,11519826 | 5,56953654 | 4,42 | 5,32 |
| II K  | 5 | 3 | 12 | 1,11519826 | 5,56953654 | 6,05 | 5,32 |
| I I   | 6 | 3 | 16 | 1,05707983 | 5,85292061 | 4,62 | 5,29 |
| I S   | 4 | 5 | 2  | 0,9520089  | 4,15888308 | 4,7  | 3,95 |



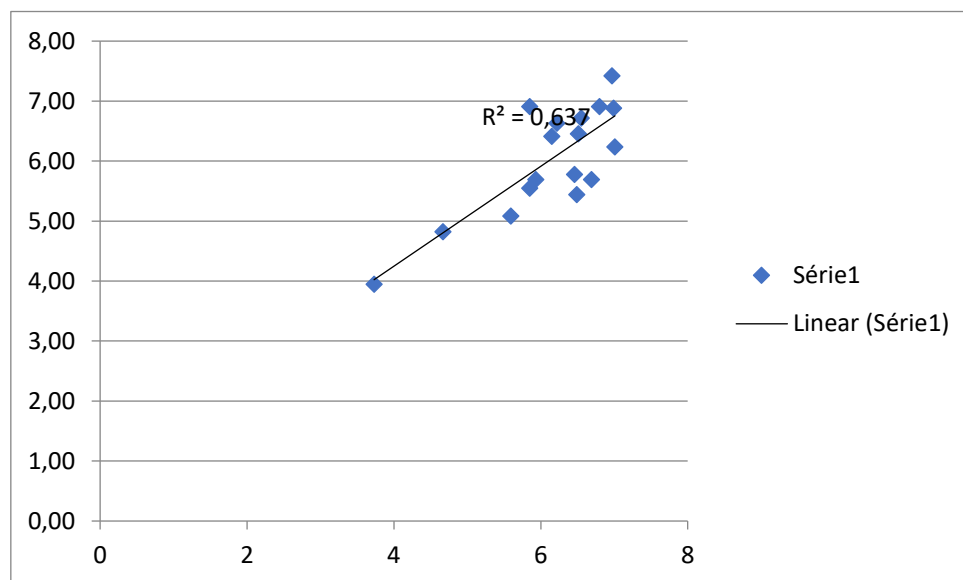
**Figura S3.** Previsão da atividade biológica com o modelo 2D com coeficiente de determinação  $R^2$

### 3.2. Teste

**Tabela S5.** Valores dos descritores do modelo 2D para o conjunto de teste

| Code | topoRadius | nHBint5 | nBondsM | GATS4c     | piPC5      | Y.EXP | y-pred |
|------|------------|---------|---------|------------|------------|-------|--------|
| II A | 4          | 3       | 3       | 1,08192133 | 4,09434456 | 4,67  | 4,82   |
| II H | 5          | 3       | 8       | 1,10795951 | 4,92316927 | 5,93  | 5,69   |
| I G  | 4          | 5       | 2       | 0,9520089  | 4,15888308 | 3,73  | 3,95   |
| IV H | 7          | 4       | 22      | 1,30371569 | 6,25670927 | 7,01  | 6,24   |
| I L  | 6          | 3       | 9       | 1,14091876 | 5,39957573 | 6,97  | 7,42   |
| IV M | 6          | 3       | 13      | 1,11805535 | 5,73778124 | 6,51  | 6,46   |
| I O  | 5          | 3       | 6       | 1,042932   | 5,24174702 | 6,15  | 6,41   |
| I N  | 6          | 3       | 9       | 1,01829303 | 5,51468803 | 6,99  | 6,88   |
| II G | 4          | 3       | 6       | 1,08049029 | 5,00478488 | 5,59  | 5,08   |

|       |   |   |    |            |            |      |      |
|-------|---|---|----|------------|------------|------|------|
| III H | 5 | 3 | 8  | 1,09472552 | 5,03897977 | 6,46 | 5,78 |
| II D  | 6 | 3 | 8  | 1,07021868 | 5,08179249 | 6,8  | 6,91 |
| II E  | 6 | 3 | 8  | 1,07021868 | 5,08179249 | 5,85 | 6,91 |
| III I | 4 | 3 | 6  | 1,12634464 | 5,07751482 | 6,49 | 5,44 |
| IV K  | 5 | 3 | 8  | 1,10795951 | 4,92316927 | 6,69 | 5,69 |
| IV L  | 6 | 3 | 12 | 1,11660407 | 5,69897929 | 6,55 | 6,72 |
| III A | 5 | 3 | 6  | 1,10306946 | 5,15256879 | 6,22 | 6,63 |
| II J  | 5 | 4 | 12 | 1,31770083 | 5,09834087 | 5,85 | 5,55 |



**Figura S4.** Previsão da atividade biológica com o modelo 2D com coeficiente de determinação  $R^2$  para o conjunto teste

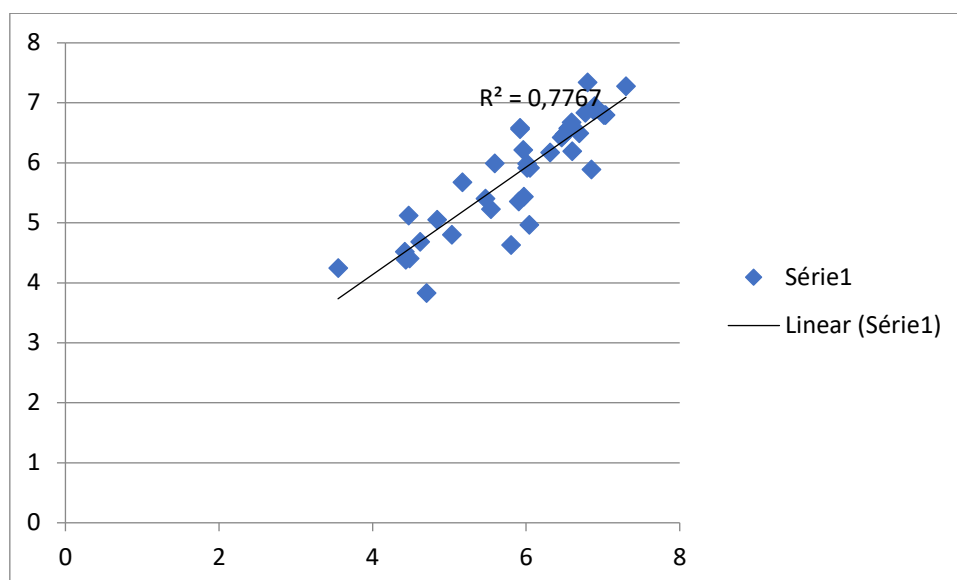
#### 4. Modelo 3D

##### 4.1. Treinamento

**Tabela S6.** Valores dos descritores 3D do modelo 2D

|       | HF_23.9645<br>0_29.94940<br>_41.78190 | QQ_24.9645<br>0_23.94940_<br>39.78190 | LJ_24.96450<br>_22.94940_4<br>7.78190 | HF_28.96450_<br>27.94940_38.7<br>8190 | HB_27.96450_<br>28.94940_41.7<br>8190 | LJ_24.96450_30<br>.94940_38.7819<br>0 | HB_23.96450_<br>28.94940_43.7<br>8190 | HF_25.96450<br>_26.94940_4<br>0.78190 | Y    | Y_PRED   |
|-------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|------|----------|
| I B   | -0,12913                              | 1,77055                               | -0,01786                              | 0,01639                               | 1,99342                               | 0,00012                               | 31,71757                              | 33,78941                              | 6,31 | 6,176152 |
| I A   | -0,13549                              | 1,13153                               | -0,0187                               | 0,02859                               | 0,65085                               | 0                                     | 2,36971                               | 15,97302                              | 5,92 | 6,562365 |
| I I   | -0,12111                              | 1,24487                               | -0,02112                              | -0,09574                              | 0,35747                               | 0,99995                               | -0,05802                              | 15,99796                              | 4,62 | 4,68469  |
| I H   | -0,09823                              | 1,58306                               | -0,014                                | 0,01584                               | 0,36875                               | 0,98999                               | 10,03182                              | 32,29964                              | 5,54 | 5,230162 |
| II I  | 32,06516                              | 1,53514                               | 10,96867                              | 0,0023                                | 0                                     | 0,02197                               | 26,74634                              | -0,0023                               | 6,01 | 5,921079 |
| II K  | 32,32506                              | 1,14408                               | 0,50457                               | -0,07231                              | 0                                     | 0,83939                               | 32,27061                              | -0,00235                              | 6,05 | 5,917008 |
| III B | -0,04094                              | 1,3179                                | -0,15418                              | 0,00491                               | 0,00537                               | 0                                     | 32,87012                              | -0,00324                              | 4,47 | 5,121816 |
| III C | -0,02468                              | 1,4356                                | 1,87546                               | 0,00088                               | 0,00693                               | 0,89108                               | 0,16873                               | -0,00347                              | 5,8  | 4,632142 |
| II G  | 31,30991                              | 1,59043                               | -0,03447                              | 0,0004                                | 0                                     | 0,64318                               | 36,5826                               | -0,00421                              | 5,59 | 5,994393 |
| III H | 31,50093                              | 1,45399                               | 3,24327                               | 0,00074                               | 0                                     | 0,00801                               | 36,17679                              | -0,00398                              | 6,46 | 6,425697 |
| III E | -0,00893                              | 1,41833                               | -0,12336                              | -0,01388                              | 0                                     | 0,00005                               | -0,3725                               | -0,00962                              | 5,97 | 5,438023 |
| III F | 35,22231                              | 1,50336                               | 31,91807                              | 0,02145                               | 0                                     | 0                                     | 34,19278                              | -0,0039                               | 4,42 | 4,520552 |
| I P   | 35,08045                              | 1,5542                                | -0,03154                              | -0,05165                              | 0                                     | 0                                     | 32,87041                              | -0,00202                              | 6,5  | 6,478081 |
| I Q   | 34,26719                              | 1,60559                               | -0,02883                              | 0,01039                               | 0                                     | 0,00001                               | 32,92907                              | -0,0021                               | 7,03 | 6,800991 |
| I R   | 35,02946                              | 1,58165                               | -0,04043                              | -0,03205                              | 0                                     | 0                                     | 32,62169                              | -0,00214                              | 5,92 | 6,584777 |
| I S   | -0,01955                              | 1,6861                                | -0,03943                              | -0,00571                              | 0                                     | 0,99074                               | 32,2948                               | -0,00296                              | 4,7  | 3,834023 |
| III D | 0,3168                                | 1,11113                               | -0,31847                              | -0,00038                              | 0,00628                               | 0                                     | 32,17734                              | -0,00367                              | 5,9  | 5,360076 |
| IV O  | -0,2796                               | 1,1395                                | -0,21743                              | -0,00189                              | 0,00002                               | 0,39592                               | 31,21552                              | -0,00342                              | 6,04 | 4,967466 |
| I T   | 1,46367                               | 1,10819                               | -0,24276                              | -0,00298                              | 0                                     | 0,74731                               | 31,68628                              | -0,00359                              | 5,03 | 4,803106 |
| II D  | 34,22932                              | 1,09714                               | -0,20102                              | 0,00244                               | 0,00527                               | 0                                     | 31,73                                 | -0,00306                              | 6,8  | 7,343603 |
| III G | 3,74129                               | 1,10908                               | 6,74733                               | -0,00154                              | 0,97076                               | 0                                     | 32,11248                              | -0,00422                              | 5,47 | 5,404457 |
| I K   | 34,53059                              | 1,64226                               | -0,04876                              | -0,07132                              | 0                                     | 0,01603                               | 33,26744                              | -0,00199                              | 6,6  | 6,194782 |
| I M   | 34,79608                              | 1,4649                                | -0,05085                              | 0,01967                               | 0,00006                               | 0,95156                               | 36,30372                              | 30,66106                              | 7,01 | 6,800161 |
| IV J  | 34,992                                | 1,5729                                | -0,04719                              | 0,02122                               | 0                                     | 0                                     | 37,66048                              | -0,00214                              | 6,88 | 6,875339 |
| IV G  | 34,96114                              | 1,61868                               | -0,04439                              | 0,03723                               | 0                                     | 0,00022                               | 36,62574                              | -0,00221                              | 6,89 | 6,950261 |
| IV M  | 34,8611                               | 1,58302                               | -0,04492                              | 0,05443                               | 0                                     | 0,74567                               | 36,71447                              | -0,00224                              | 6,69 | 6,49476  |
| II B  | 33,05098                              | 1,34922                               | -0,05509                              | -0,03572                              | 0                                     | 0,79352                               | 36,6639                               | -0,00205                              | 6,01 | 5,990835 |
| IV I  | 34,30691                              | 1,50666                               | -0,04856                              | -0,00252                              | 0,00009                               | 0,1093                                | 36,19633                              | 31,6979                               | 7,3  | 7,279086 |
| IV E  | 34,58324                              | 1,75859                               | -0,05213                              | -0,08385                              | 0,0004                                | 0,65935                               | 37,20788                              | 15,10549                              | 5,17 | 5,677865 |
| IV P  | 32,97794                              | 1,50804                               | -0,05174                              | -0,05265                              | 0,00001                               | 0,54333                               | 37,14811                              | 0,24372                               | 6,85 | 5,893545 |
| IV B  | 33,05487                              | 0,81451                               | -0,04348                              | -0,20517                              | 0,00598                               | 0,073                                 | 35,58681                              | 34,64871                              | 6,59 | 6,674324 |
| IV C  | 34,4294                               | 1,69153                               | -0,05449                              | -0,18352                              | 0,00003                               | 0,1836                                | 42,90179                              | -0,28042                              | 4,84 | 5,057069 |

|      |          |         |          |          |         |         |          |          |      |          |
|------|----------|---------|----------|----------|---------|---------|----------|----------|------|----------|
| IV A | 34,39938 | 1,64136 | -0,14772 | 0,0154   | 0,00002 | 0,64967 | 36,92499 | -0,00426 | 5,96 | 6,216919 |
| IV L | 35,02311 | 1,80466 | -0,04968 | 0,01254  | 0,00003 | 0       | 36,34635 | -0,16863 | 6,55 | 6,579528 |
| IV D | 33,73426 | 1,63775 | -0,04874 | -0,07457 | 0,00046 | 0       | 37,84523 | 26,69344 | 6,55 | 6,570869 |
| I F  | -0,01108 | 1,6714  | -0,01106 | -0,00581 | 0       | 0,53249 | 30,62673 | -0,00191 | 3,55 | 4,248826 |
| I E  | -0,0247  | 2,00903 | -0,0227  | -0,00003 | 0,02182 | 0,00504 | 32,70487 | 3,78567  | 4,43 | 4,389153 |
| I J  | -0,01349 | 1,94074 | -0,02113 | 0,00039  | 0,00482 | 0,00021 | 31,48581 | -0,21444 | 4,48 | 4,4102   |
| II C | 34,69544 | 1,60708 | -0,04629 | 0,01232  | 0       | 0,00002 | 33,11368 | -0,00208 | 6,77 | 6,835533 |

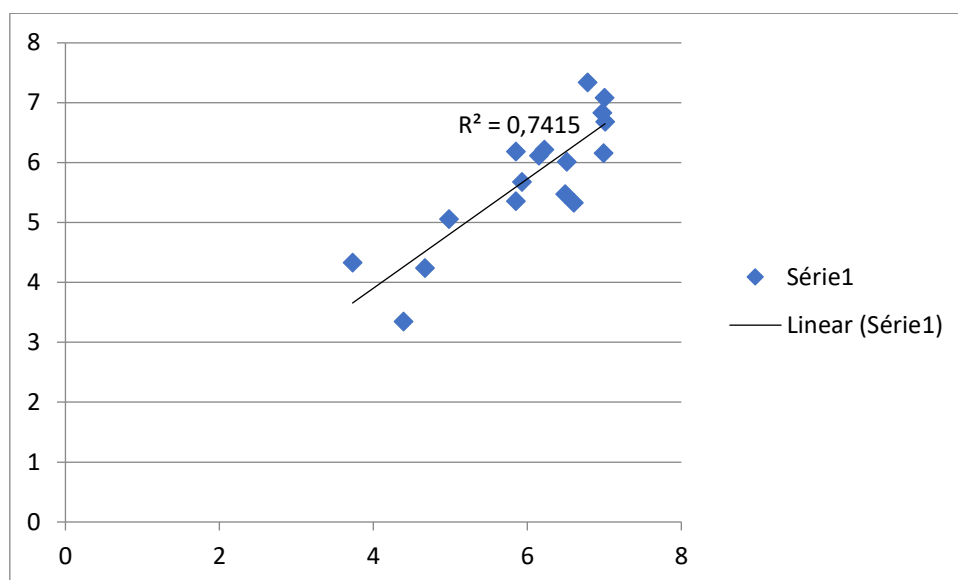


**Figura S5.** Previsão da atividade biológica com o modelo 3D com coeficiente de determinação  $R^2$

#### 4.2. Teste

**Tabela S7.** Valores dos descritores do modelo 3D para o conjunto de teste

|       | HF_23.9645<br>0_29.94940 | QQ_24.9645<br>0_23.94940_ | LJ_24.96450_2<br>2.94940_47.78 | HF_28.96450<br>_27.94940_3 | HB_27.96450_<br>28.94940_41.7 | LJ_24.96450<br>_30.94940_3 | HB_23.96450_<br>28.94940_43.7 | HF_25.96450_<br>26.94940_40.7 |      |          |
|-------|--------------------------|---------------------------|--------------------------------|----------------------------|-------------------------------|----------------------------|-------------------------------|-------------------------------|------|----------|
| Code  | _41.78190                | 39.78190                  | 190                            | 8.78190                    | 8190                          | 8.78190                    | 8190                          | 8190                          | Y    | Y_PRED   |
| I C   | -0,05872                 | 1,88618                   | -0,04094                       | -0,10778                   | 0,05255                       | 0,40655                    | 35,5913                       | 0,19496                       | 4,39 | 3,351666 |
| II A  | -0,04563                 | 1,39539                   | -0,22481                       | -0,11326                   | 0                             | 0,00211                    | 31,68579                      | -0,00237                      | 4,67 | 4,242175 |
| II H  | 27,6399                  | 1,48214                   | 13,98838                       | 0,00275                    | 0                             | 0                          | 16,70699                      | -0,00218                      | 5,93 | 5,678786 |
| I G   | -0,06809                 | 1,83775                   | -0,02344                       | -0,01089                   | 0                             | 0,03681                    | 36,37585                      | -0,00233                      | 3,73 | 4,333426 |
| IV H  | 34,61941                 | 1,65803                   | -0,06374                       | -0,0007                    | 0                             | 0                          | 33,30198                      | -0,00207                      | 7,01 | 6,683087 |
| I L   | 34,69544                 | 1,60708                   | -0,04629                       | 0,01232                    | 0                             | 0,00002                    | 33,11368                      | -0,00208                      | 6,97 | 6,835533 |
| IV M  | 33,05193                 | 1,53767                   | -0,05355                       | -0,09526                   | 0                             | 0,00003                    | 36,66744                      | -0,00221                      | 6,51 | 6,016677 |
| I O   | 34,73298                 | 1,56354                   | -0,03738                       | -0,05613                   | 0                             | 0,37113                    | 32,66828                      | -0,00199                      | 6,15 | 6,118777 |
| I N   | 34,64896                 | 1,60568                   | -0,05512                       | -0,01749                   | 0                             | 0,57458                    | 33,1528                       | -0,00204                      | 6,99 | 6,161425 |
| II J  | -0,0658                  | 1,11046                   | -0,15258                       | 0,00418                    | 0                             | 0                          | 31,81402                      | -0,00268                      | 5,85 | 5,360759 |
| II F  | -0,2685                  | 1,08313                   | -0,20207                       | -0,00227                   | 0                             | 0                          | 32,03667                      | -0,0028                       | 6,6  | 5,33504  |
| II E  | 31,94943                 | 1,47541                   | 6,15998                        | 0,0528                     | 0                             | 0,94754                    | 12,65132                      | -0,0022                       | 5,85 | 6,190302 |
| III I | 2,93106                  | 1,0738                    | 0,10246                        | -0,00187                   | 0,00004                       | 0,03801                    | 32,01474                      | -0,00275                      | 6,49 | 5,47778  |
| III A | -0,0195                  | 1,15068                   | -0,07674                       | -0,00889                   | 1,82833                       | 0,9947                     | 6,16919                       | 32,05432                      | 6,22 | 6,219159 |
| IV F  | 0                        | 0                         | 0                              | -0,00072                   | 0                             | 0                          | 0                             | 0                             | 7    | 7,086064 |
| IV N  | 34,22932                 | 1,09714                   | -0,20102                       | 0,00244                    | 0,00527                       | 0                          | 31,73                         | -0,00306                      | 6,78 | 7,3406   |
| I D   | -0,01329                 | 1,94543                   | -0,02186                       | 0,01728                    | 0,00855                       | 0,00005                    | 0,01812                       | 0,13001                       | 4,98 | 5,059978 |



**Figura S6.** Previsão da atividade biológica com o modelo 3D com coeficiente de determinação  $R^2$  para o conjunto teste