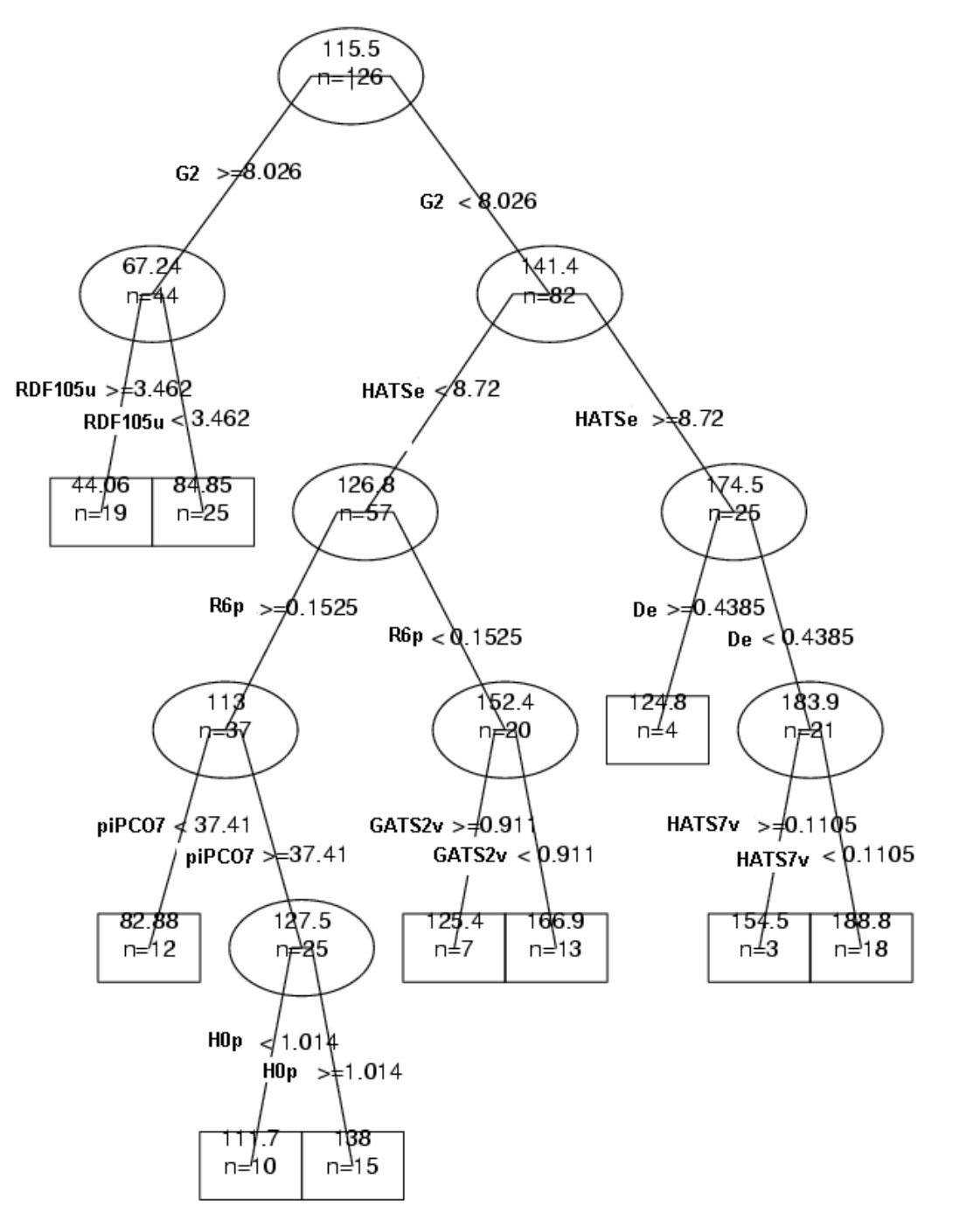
Supplementary Material (ESI) for Green Chemistry

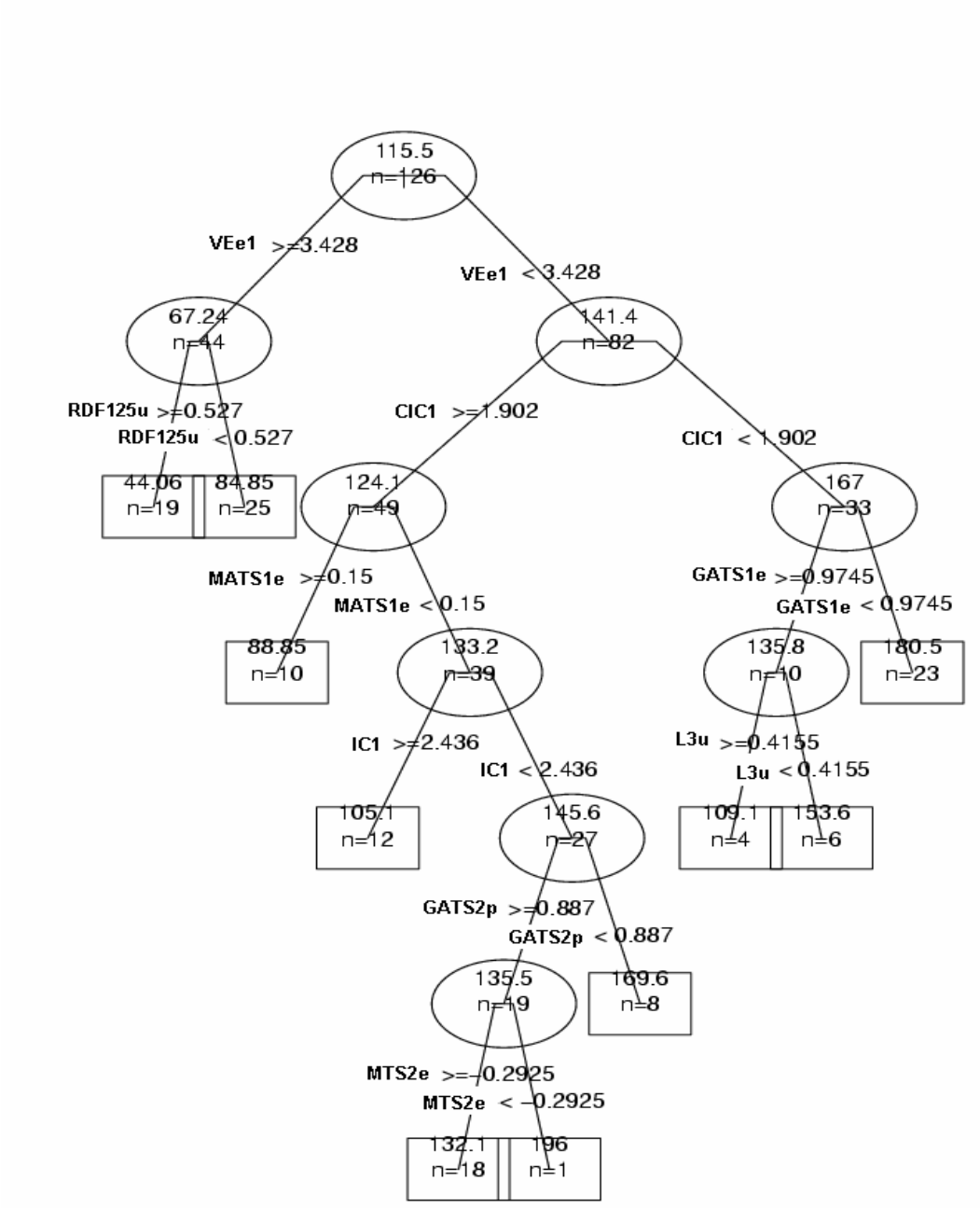
Electronic Supplementary Information

I) Schemes of the 10 trees used in figure 5 of the article

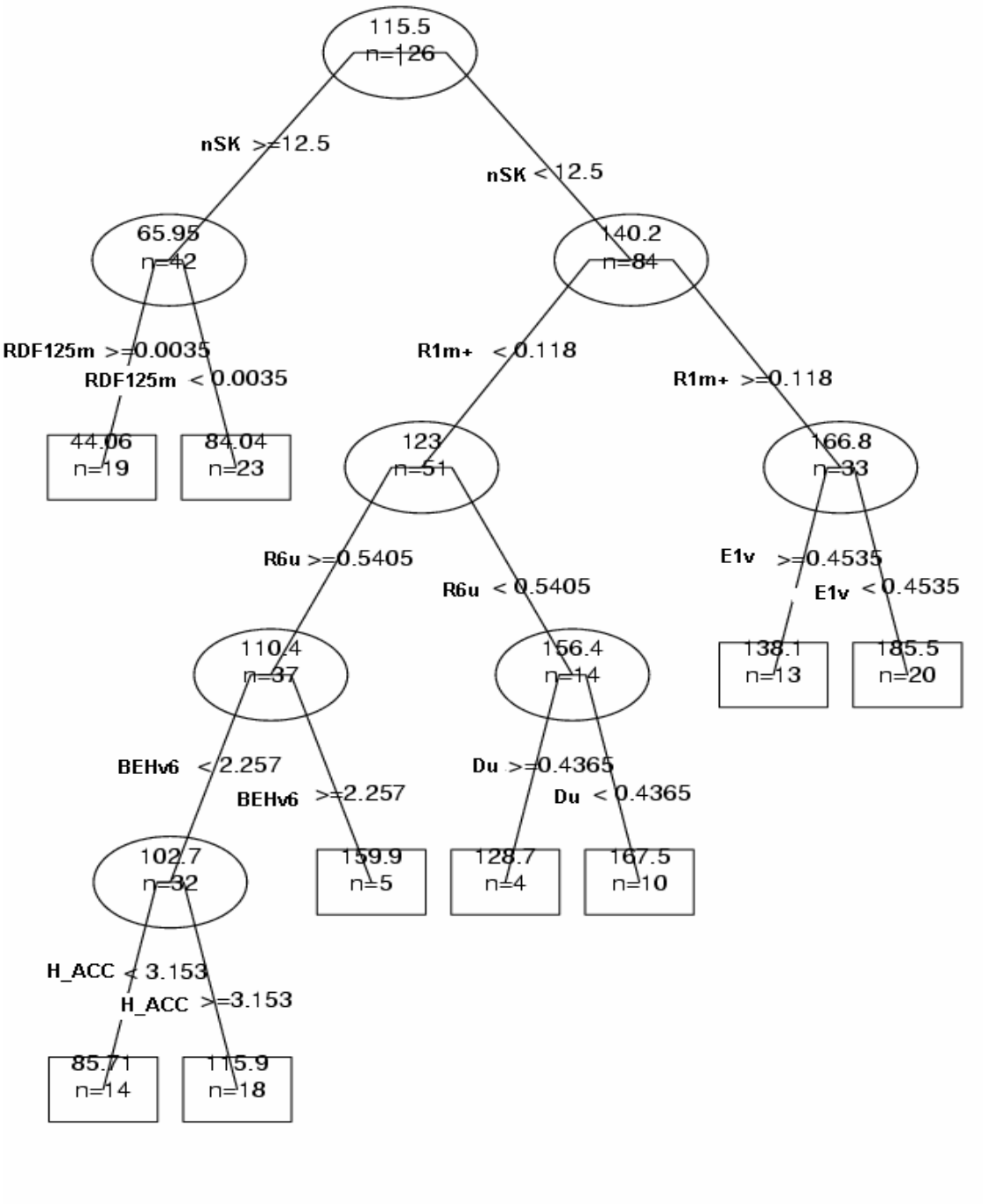
Tree 1



Tree 2

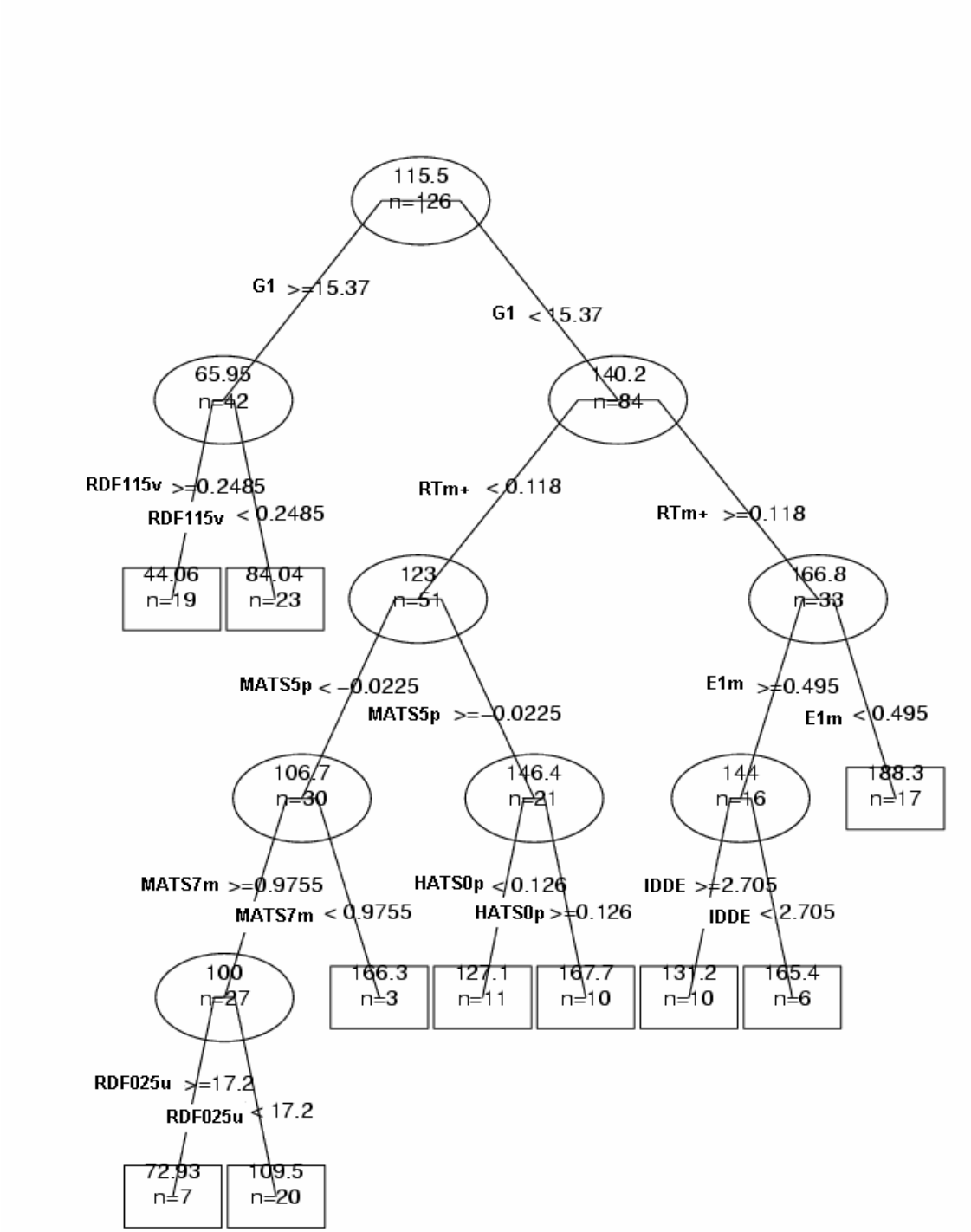


Tree 3

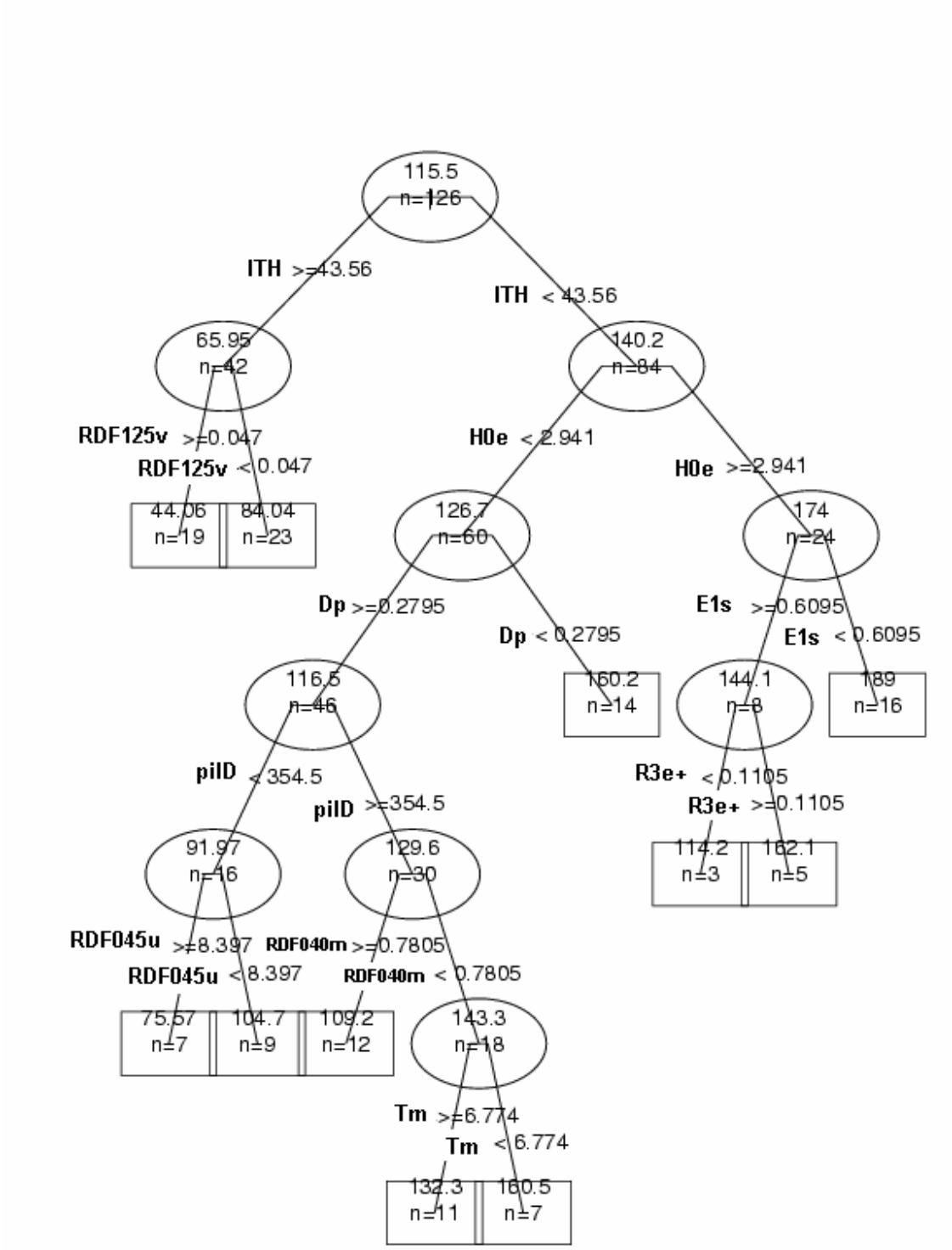


Note: H_ACC (in electron units) is the highest H acceptor potential as calculated by PETRA software.

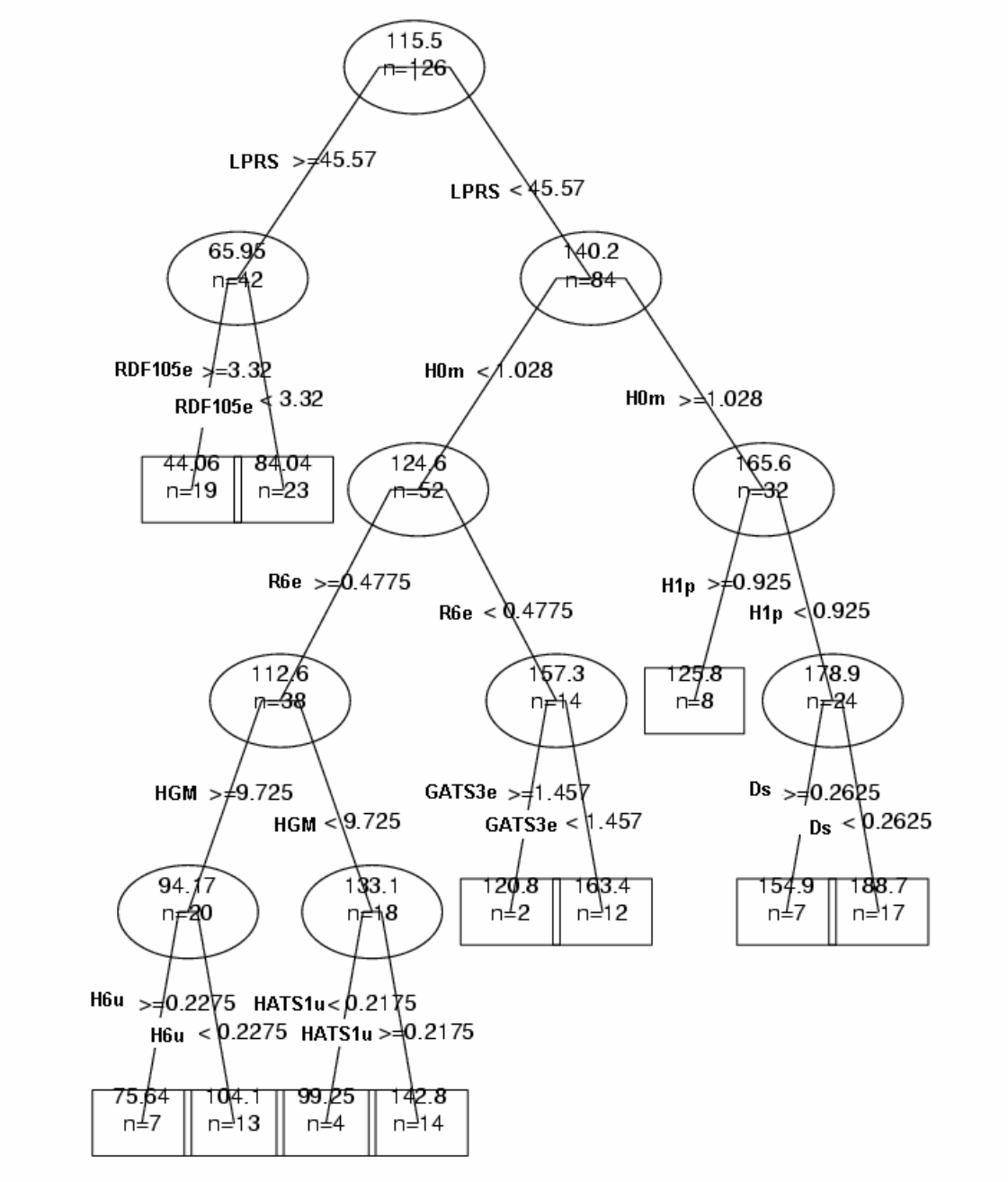
Tree 4



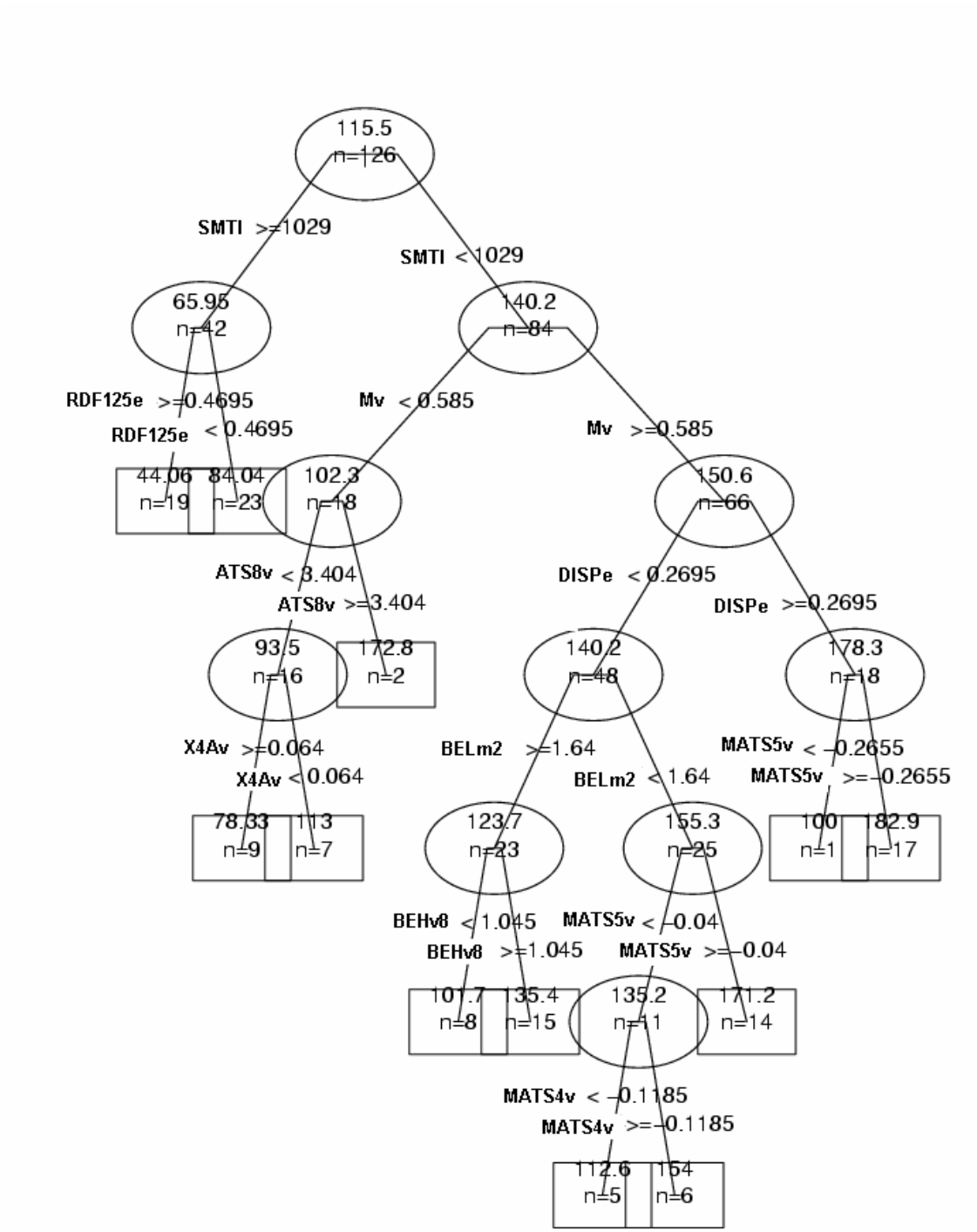
Tree 5



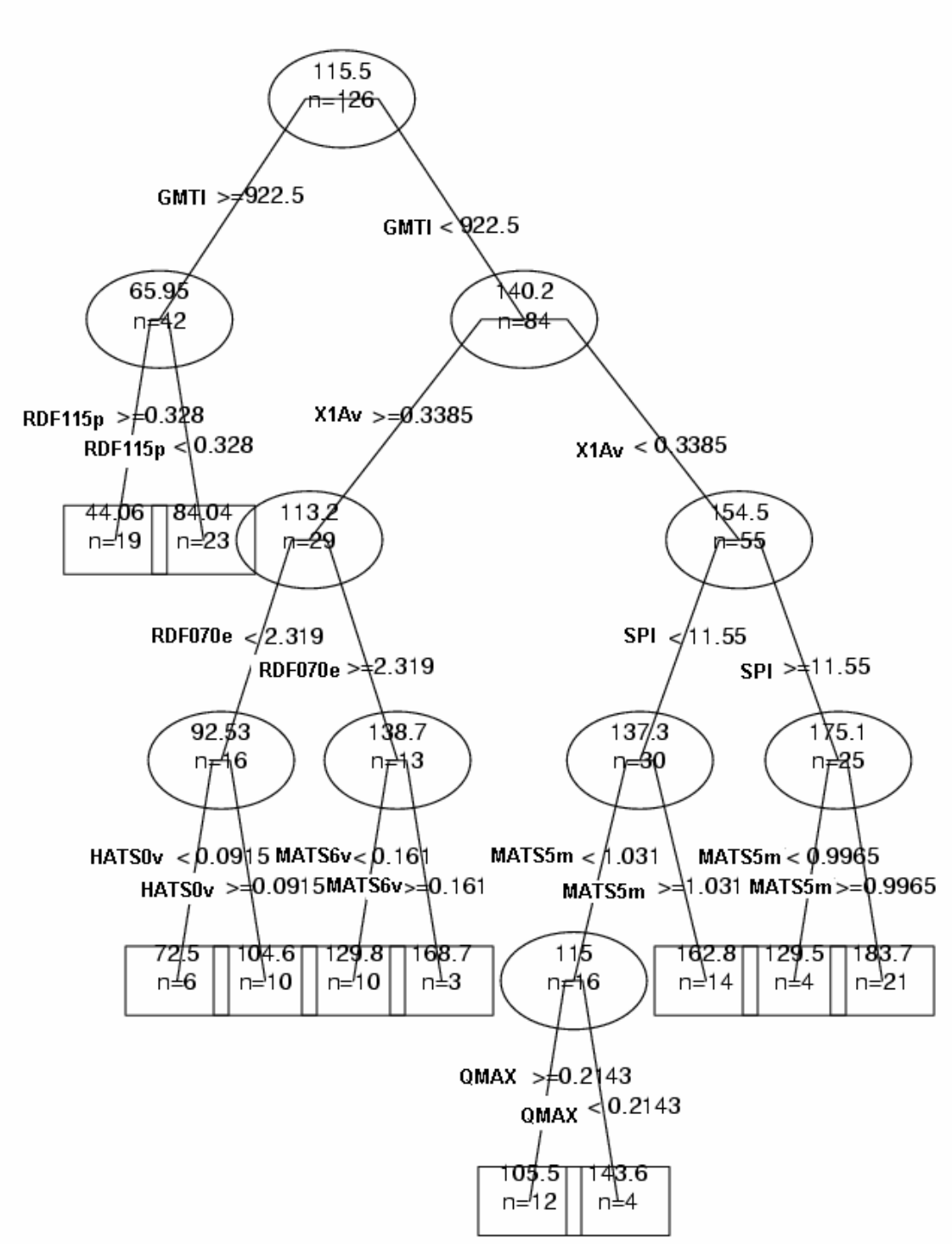
Tree 6



Tree 7

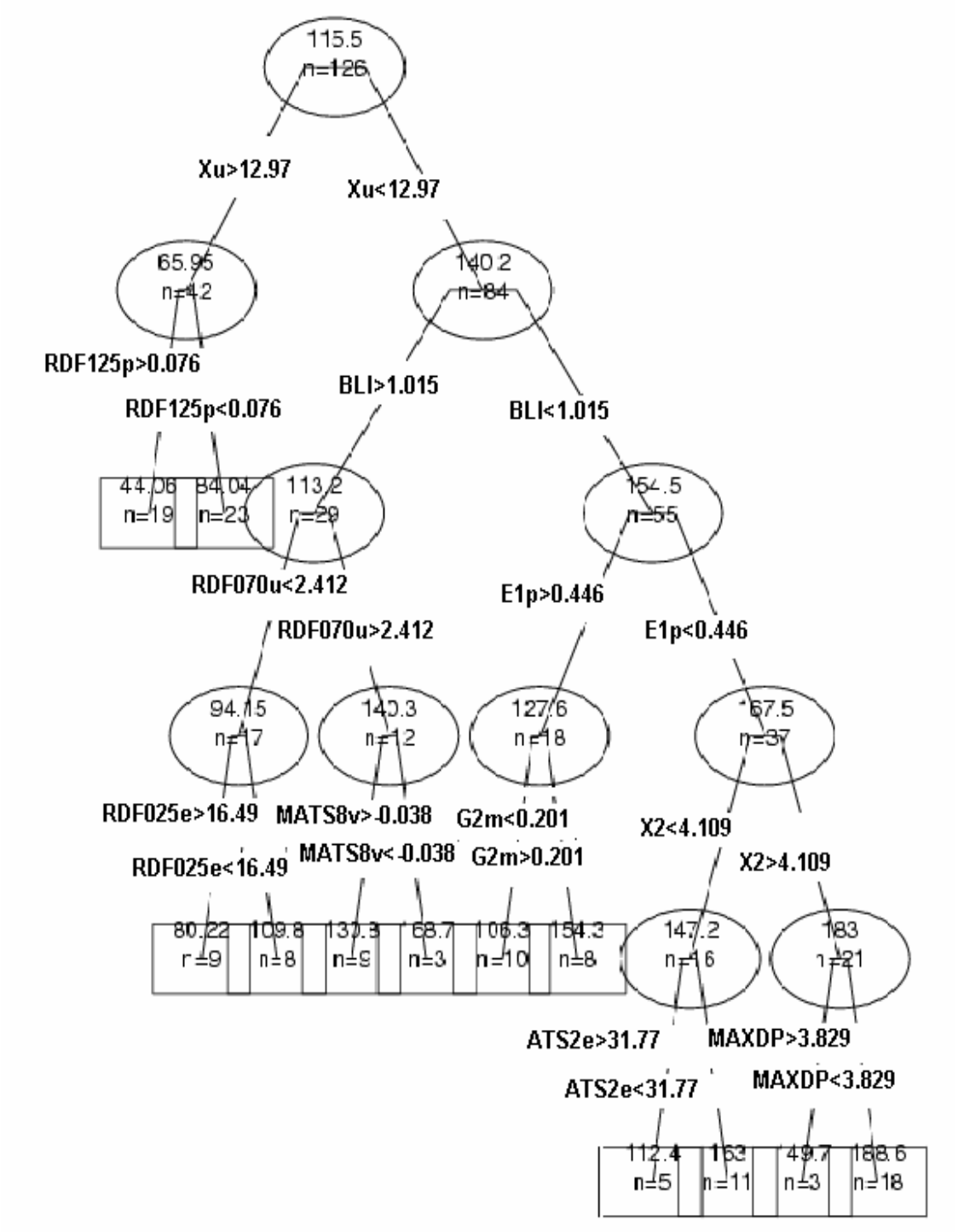


Tree 8

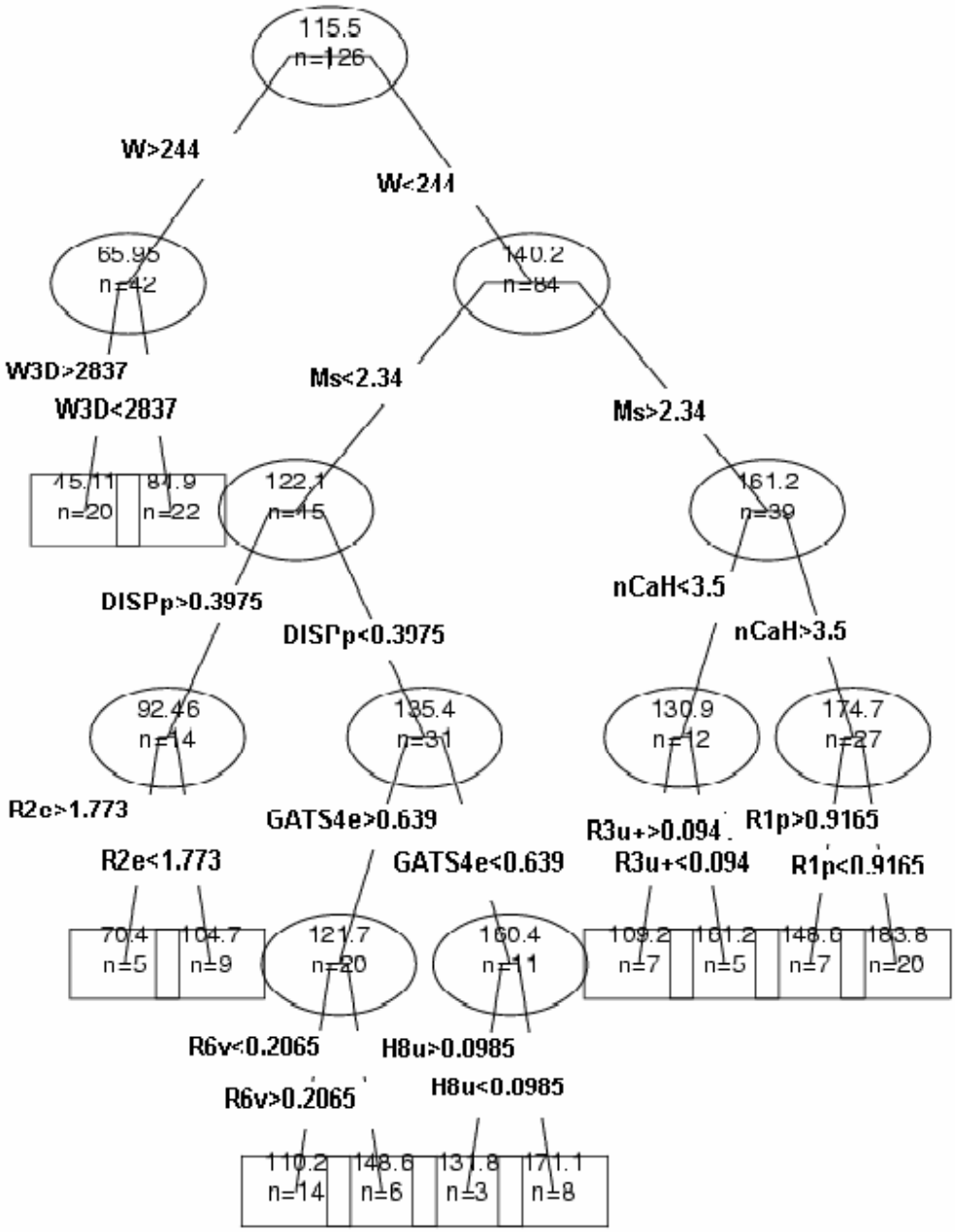


Note: QMAX (in electron units) is the highest partial atomic charge in a molecule as calculated by PETRA 3.1 software.

Tree 9



Tree 10



II) List of all the descriptors used in the initial pool of descriptors

2D-Autocorrelation - Moreau-Broto (ATS), Moran (MATS) and Geary (GATS) autocorrelations, by summing the products of all terminal atoms of a considered molecular path			
ATS1m	ATS7e	MATS6v	GATS2v
ATS2m	ATS8e	MATS7v	GATS3v
ATS3m	ATS1p	MATS8v	GATS4v
ATS4m	ATS2p	MATS1e	GATS5v
ATS5m	ATS3p	MATS2e	GATS6v
ATS6m	ATS4p	MATS3e	GATS7v
ATS7m	ATS5p	MATS4e	GATS8v
ATS8m	ATS6p	MATS5e	GATS1e
ATS1v	ATS7p	MATS6e	GATS2e
ATS2v	ATS8p	MATS7e	GATS3e
ATS3v	MATS1m	MATS8e	GATS4e
ATS4v	MATS2m	MATS1p	GATS5e
ATS5v	MATS3m	MATS2p	GATS6e
ATS6v	MATS4m	MATS3p	GATS7e
ATS7v	MATS5m	MATS4p	GATS8e
ATS8v	MATS6m	MATS5p	GATS1p
ATS1e	MATS7m	MATS6p	GATS2p
ATS2e	MATS8m	MATS7p	GATS3p
ATS3e	MATS1v	MATS8p	GATS4p
ATS4e	MATS2v	GATS6m	GATS5p
ATS5e	MATS3v	GATS7m	GATS6p
ATS6e	MATS4v	GATS8m	GATS7p
ATS1m	MATS5v	GATS1v	GATS8p

WHIM – Descriptors obtained projecting the atomic coordinates in the three main axis				
L1u	L2v	L3p	Te	Ds
L2u	L3v	P1p	Tp	Vu
L3u	P1v	P2p	Ts	Vm
P1u	P2v	G1p	Au	Vv
P2u	G1v	G2p	Am	Ve
G1u	G2v	G3p	Av	
G2u	G3v	E1p	Ae	
G3u	E1v	E2p	Ap	
E1u	E2v	E3p	As	
E2u	E3v	L1s	Gu	
E3u	L1e	L2s	Gm	
L1m	L2e	L3s	Gs	
L2m	L3e	P1s	Ku	
L3m	P1e	P2s	Km	
P1m	P2e	G1s	Kv	
P2m	G1e	G2s	Ke	
G1m	G2e	G3s	Kp	
G2m	G3e	E1s	Ks	
G3m	E1e	E2s	Du	

E1m	E2e	E3s	Dm	
E2m	E3e	Tu	Dv	
E3m	L1p	Tm	De	
L1v	L2p	Tv	Dp	

AROMATIC – Aromaticity indexes
HOMA
RCI
AROM
HOMT

ATOMCENT- accounts for the number of substructures centered in a specific atom (C, N, O…)	
C-001	H-047
C-002	H-048
C-003	H-049
C-005	H-050
C-006	H-051
C-008	H-052
C-009	H-053
C-015	H-054
C-016	O-056
C-017	O-057
C-018	O-058
C-021	O-060
C-022	O-061
C-024	N-072
C-025	N-074
C-026	N-075
C-027	F-081
C-028	
C-031	
C-038	
C-039	
C-040	
H-046	

BCUT- Molecular descriptors based on the eigenvalues extracted from Burden adjacency matrix		
BEHm1	BEHv8	BELe7
BEHm2	BELv1	BELe8
BEHm3	BELv2	BEHp1
BEHm4	BELv3	BEHp2
BEHm5	BELv4	BEHp3
BEHm6	BELv5	BEHp4
BEHm7	BELv6	BEHp5
BEHm8	BELv7	BEHp6
BELm1	BELv8	BEHp7
BELm2	BEHe1	BEHp8
BELm3	BEHe2	BELp1
BELm4	BEHe3	BELp2

BELm5	BEHe4	BELp3
BELm6	BEHe5	BELp4
BELm7	BEHe6	BELp5
BELm8	BEHe7	BELp6
BEHv1	BEHe8	BELp7
BEHv2	BELe1	BELp8
BEHv3	BELe2	
BEHv4	BELe3	
BEHv5	BELe4	
BEHv6	BELe5	
BEHv7	BELe6	

CONSTITUTIONAL- Global characteristics of a molecule like molecular weight or the number of atoms, not accounting for the connectivity or conformation of the molecule.	
MW	nH
AMW	nC
Sv	nN
Se	nO
Sp	nF
Ss	nX
Mv	nR06
Me	nBnz
Mp	
Ms	
nAT	
nSK	
nBT	
nBO	
nBM	
SCBO	
nCIC	
nCIR	
RBN	
RBF	
nDB	
nTB	
nAB	

EMPIRICAL – Molecular descriptors based on the number of non-single bonds (Ui), number f hydrophilic groups (Hy), and ratio between the aromatic bond and other bonds in a H depleted molecule (ARR)
Ui
Hy
ARR

FUNCTIONAL GROUP – Molecular descriptors based on the number of certain functionalities in a molecule	
nCp	nOHp
nCs	nRORPh

nCt	nRCH2X
nCrH2	nHDon
nCrHR	nHAcc
nCaH	
nCaR	
nCconjR	
n=CH2	
n=CHR	
n=CR2	
n#CH	
n#CR	
nCOOH	
nCOOR	
nCOORPh	
nCONH2	
nCONH2Ph	
nCONR2Ph	
nCO	
nCN	
nOH	
nOHPh	

GALVEZ – Molecular descriptors based on the 10 first eigenvalues obtained from the adjacency matrix.
GGI1
GGI2
GGI3
GGI4
GGI5
GGI6
GGI7
GGI8
GGI9
GGI10
JGI1
JGI2
JGI3
JGI4
JGI5
JGI6
JGI7
JGT

GEOMETRIC – Conformationally dependent descriptors based on the molecular geometry	
W3D	QXXv
J3D	QYYv
H3D	QZZv

AGDD	DISPe
DDI	QXXe
ADDD	QYYe
G1	QZZe
G2	DISPp
RGyr	QXXp
SPAN	QYYp
SPAM	QZZp
MEcc	G(N..N)
SPH	G(N..O)
ASP	G(N..F)
FDI	G(O..O)
PJI3	
L/Bw	
SEig	
DISPm	
QXXm	
QYYm	
QZZm	
DISPv	

GETAWAY – Descriptors based on the elements of the leverage matrix obtained by the centered atomic coordinates								
ITH	H1m	H4v	H7e	HATS0p	R2u+	R7m+	R3e	R8p
ISH	H2m	H5v	H8e	HATS1p	R3u+	R8m+	R4e	RTp
HIC	H3m	H6v	HTe	HATS2p	R4u+	RTm+	R5e	R1p+
HGM	H4m	H7v	HATS0e	HATS3p	R5u+	R1v	R6e	R2p+
H1u	H5m	H8v	HATS1e	HATS4p	R6u+	R2v	R7e	R3p+
H2u	H6m	HTv	HATS2e	HATS5p	R7u+	R3v	R8e	R4p+
H3u	H7m	HATS0v	HATS3e	HATS6p	R8u+	R4v	RTe	R5p+
H4u	H8m	HATS1v	HATS4e	HATS7p	RTu+	R5v	R1e+	R6p+
H5u	HTm	HATS2v	HATS5e	HATS8p	R1m	R6v	R2e+	R7p+
H6u	HATS0m	HATS3v	HATS6e	HATSp	R2m	R7v	R3e+	R8p+
H7u	HATS1m	HATS4v	HATS7e	RCON	R3m	R8v	R4e+	RTp+
H8u	HATS2m	HATS5v	HATS8e	RARS	R4m	RTv	R5e+	
HTu	HATS3m	HATS6v	HATSe	REIG	R5m	R1v+	R6e+	
HATS0u	HATS4m	HATS7v	H0p	R1u	R6m	R2v+	R7e+	
HATS1u	HATS5m	HATS8v	H1p	R2u	R7m	R3v+	R8e+	
HATS2u	HATS6m	HATSv	H2p	R3u	R8m	R4v+	RTe+	
HATS3u	HATS7m	H0e	H3p	R4u	RTm	R5v+	R1p	
HATS4u	HATS8m	H1e	H4p	R5u	R1m+	R6v+	R2p	
HATS5u	HATSm	H2e	H5p	R6u	R2m+	R7v+	R3p	
HATS6u	H0v	H3e	H6p	R7u	R3m+	R8v+	R4p	
HATS7u	H1v	H4e	H7p	R8u	R4m+	RTv+	R5p	
HATS8u	H2v	H5e	H8p	RTu	R5m+	R1e	R6p	
H0m	H3v	H6e	HTp	R1u+	R6m+	R2e	R7p	

MOLWALK – Molecular descriptors obtained from molecular graphs based on counting walks of a specific length
MWC01
MWC02
MWC03
MWC04
MWC05
MWC06
MWC07
MWC08
MWC09
TWC
SRW02
SRW04
SRW06
SRW08
SRW10

PROPERTI – Molecular properties calculated by regression models
MR
PSA
MLOGP

RANDICMO- Molecular descriptors derived from the distance distribution moments of the geometry matrix	
DP01	SP04
DP02	SP05
DP03	SP06
DP04	SP07
DP05	SP08
DP06	SP09
DP07	SP10
DP08	SP11
DP09	SP12
DP10	SP13
DP11	SP14
DP12	SP15
DP13	SP16
DP14	SP17
DP15	SP18
DP16	SP19
DP17	SP20
DP18	SHP2
DP19	
DP20	
SP01	
SP02	
SP03	

RDF – molecular descriptors based in a radial function and that accounts for the number of pairs of atoms at a certain distance in Å.						
RDF010u	RDF125u	RDF090m	RDF055v	RDF020e	RDF135e	RDF100p
RDF015u	RDF130u	RDF095m	RDF060v	RDF025e	RDF140e	RDF105p
RDF020u	RDF135u	RDF100m	RDF065v	RDF030e	RDF145e	RDF110p
RDF025u	RDF140u	RDF105m	RDF070v	RDF035e	RDF150e	RDF115p
RDF030u	RDF145u	RDF110m	RDF075v	RDF040e	RDF155e	RDF120p
RDF035u	RDF150u	RDF115m	RDF080v	RDF045e	RDF010p	RDF125p
RDF040u	RDF155u	RDF120m	RDF085v	RDF050e	RDF015p	RDF130p
RDF045u	RDF010m	RDF125m	RDF090v	RDF055e	RDF020p	RDF135p
RDF050u	RDF015m	RDF130m	RDF095v	RDF060e	RDF025p	RDF140p
RDF055u	RDF020m	RDF135m	RDF100v	RDF065e	RDF030p	RDF145p
RDF060u	RDF025m	RDF140m	RDF105v	RDF070e	RDF035p	RDF150p
RDF065u	RDF030m	RDF145m	RDF110v	RDF075e	RDF040p	RDF155p
RDF070u	RDF035m	RDF150m	RDF115v	RDF080e	RDF045p	
RDF075u	RDF040m	RDF155m	RDF120v	RDF085e	RDF050p	
RDF080u	RDF045m	RDF010v	RDF125v	RDF090e	RDF055p	
RDF085u	RDF050m	RDF015v	RDF130v	RDF095e	RDF060p	
RDF090u	RDF055m	RDF020v	RDF135v	RDF100e	RDF065p	
RDF095u	RDF060m	RDF025v	RDF140v	RDF105e	RDF070p	
RDF100u	RDF065m	RDF030v	RDF145v	RDF110e	RDF075p	
RDF105u	RDF070m	RDF035v	RDF150v	RDF115e	RDF080p	
RDF110u	RDF075m	RDF040v	RDF155v	RDF120e	RDF085p	
RDF115u	RDF080m	RDF045v	RDF010e	RDF125e	RDF090p	
RDF120u	RDF085m	RDF050v	RDF015e	RDF130e	RDF095p	

TOPOLOGICAL – 2D descriptors conformationally independent									
ISIZ	SPI	Jhetv	X4v	PW2	IVDM	TIC3	SEigv	VEe1	piPC09
IAC	W	Jhete	X5v	PW3	HVcpx	SIC3	SEige	VEe2	piPC10
AAC	WA	Jhetp	X0Av	PW4	HDcpx	CIC3	SEigp	VRe1	TPC
ZM1	RDSUM	MAXDN	X1Av	PW5	Uindex	BIC3	VEA1	VRe2	piID
ZM1V	Har	MAXDP	X2Av	PJI2	Vindex	IC4	VEA2	VEp1	PCR
ZM2	QW	DELS	X3Av	CSI	Xindex	TIC4	VRA1	VEp2	PCD
ZM2V	TI1	TIE	X4Av	ECC	Yindex	SIC4	VRA2	VRp1	CID
Qindex	TI2	X0	X5Av	AECC	IC0	CIC4	VED1	VRp2	BID
SNar	HyDp	X1	X0sol	DECC	TIC0	BIC4	VED2	NGS	D/Dr06
HNar	RHyDp	X2	X1sol	MDDD	SIC0	IC5	VRD1	MPC03	T(N..N)
GNar	w	X3	X2sol	UNIP	CIC0	TIC5	VRD2	MPC04	T(N..O)
Xt	ww	X4	X3sol	CENT	BIC0	SIC5	VEZ1	MPC05	T(N..F)
Dz	Rww	X5	X4sol	VAR	IC1	CIC5	VEZ2	MPC06	T(O..O)
Ram	D/D	X0A	X5sol	BAC	TIC1	BIC5	VRZ1	MPC07	
Pol	Wap	X1A	XMOD	Lop	SIC1	LP1	VRZ2	MPC08	
LPRS	WhetZ	X2A	RDCHI	ICR	CIC1	STN	VEm1	MPC09	
VDA	Whetm	X3A	RDSQ	IDE	BIC1	Eig1Z	VEm2	MPC10	
MSD	Whetv	X4A	S0K	IDM	IC2	Eig1m	VRm1	piPC03	
SMTI	Whete	X5A	S1K	IDDE	TIC2	Eig1v	VRm2	piPC04	
SMTIV	Whetp	X0v	S2K	IDDM	SIC2	Eig1e	VEv1	piPC05	
GMTI	J	X1v	S3K	IDET	CIC2	Eig1p	VEv2	piPC06	
GMTIV	JhetZ	X2v	PHI	IDMT	BIC2	SEigZ	VRv1	piPC07	
Xu	Jhetm	X3v	BLI	IVDE	IC3	SEigm	VRv2	piPC08	