SmartNanoTox International Online Conference

24 June 2020

SmartNanoTox

Prediction of NM toxicity from intelligent quantitative structure-activity relations (iQSAR)

Dr Marc Meunier

Senior Scientific Consultant and BIOVIA Fellow

Dassault Systèmes, BIOVIA, Cambridge, U.K.

3DEXPERIENCE®

DASSAULT SYSTEMES | The 3DEXPERIENCE® Company

QNTR

AGENDA

- Introduction:
  - Dassault Systèmes & BIOVIA
  - QNTR
- Toxicity Data sets
- Nanomaterials and Atomistic models
- Types and Classes of Descriptors
- Predictive Models
- Conclusion

Dassault Systèmes

A purpose-driven company

Combining Art, Science & Technology
for a more sustainable world

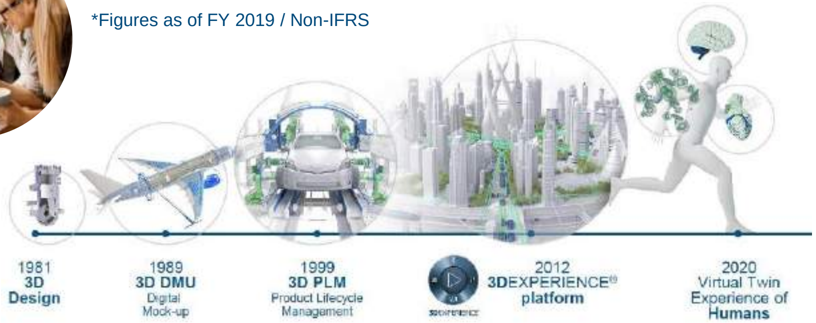
20,000 passionate people

140 nationalities
195 sites
One global R&D / 69 labs

Long-term driven

Majority shareholder control
Revenue: €4,056 millions*
Operating margin: 32%*

*Figures as of FY 2019 / Non-IFRS



12,600 partners

Software, Technology & Architecture
Content & Online services
Sales
Consulting & System Integrators (C&SI)
Education
Research

270,000 customers

11 industries in 140 countries
25 million users
Game-changing
3DEXPERIENCE platform

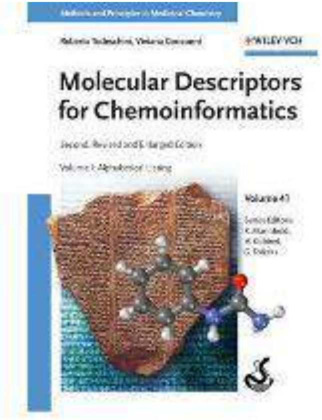
Dassault Systèmes **BIO**VIA

MODEL THE BIOSPHERE

- Focused on Supporting Science-Driven Companies
- Providing integrated, enterprise software experiences
- Delivering the breath and depth to cover the entire product life cycle

Scientific Informatics	Transform Scientific Data into Knowledge
Molecular Modeling & Simulation	Foster Innovation with *In Silico* Design
Data Science	Drive Knowledge-based Decisions
Laboratory Informatics	Optimize Lab Productivity and Compliance
Formulation Design	Accelerate Product Launches
Quality & Compliance	Drive Data-centric Quality Excellence in BioPharma
Manufacturing Analytics	Empower Production Operations in Process Industries

Quantitative Structure-Activity Relationship (QSAR)



Conditions: Y observations (Dependent variable)

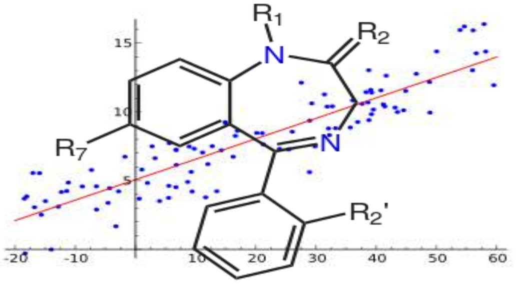
X parameters (Independent variable)

Objective: Correlate Y with X_1, X_2 ...

$$Y = f(\alpha X_1 + \beta X_2 + \cdots)$$

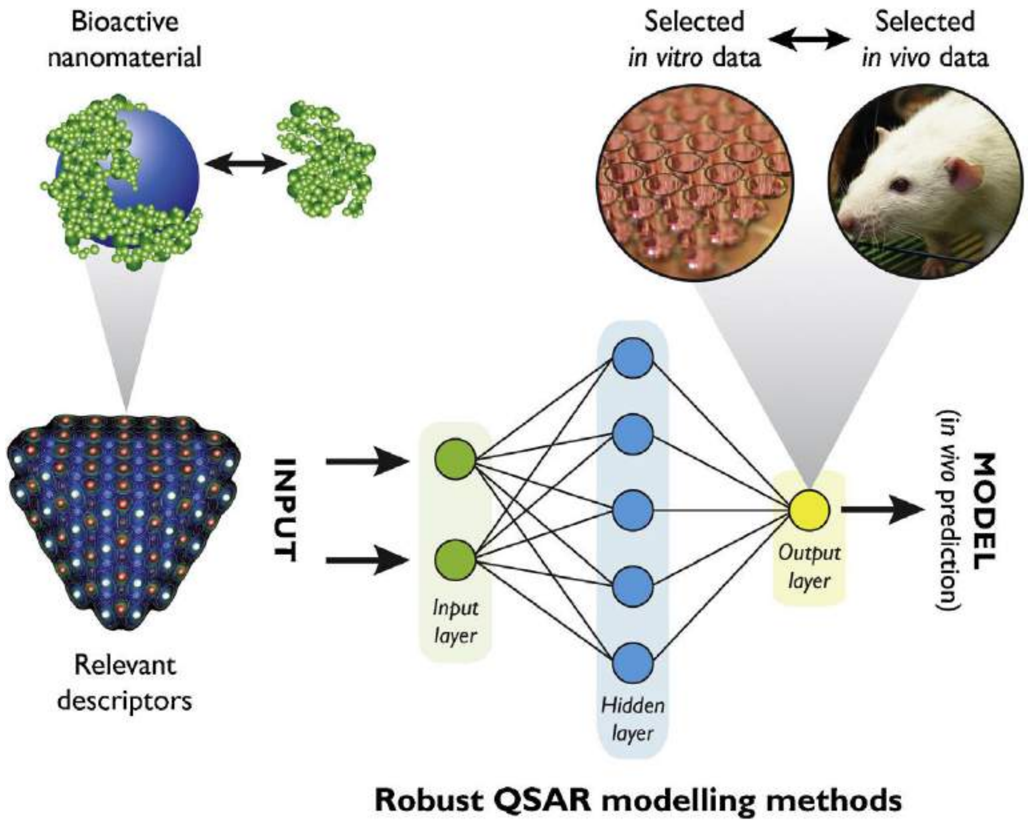
Challenge: Variance is spread over X parameters

Find the QSAR signal ... in a huge field of variance



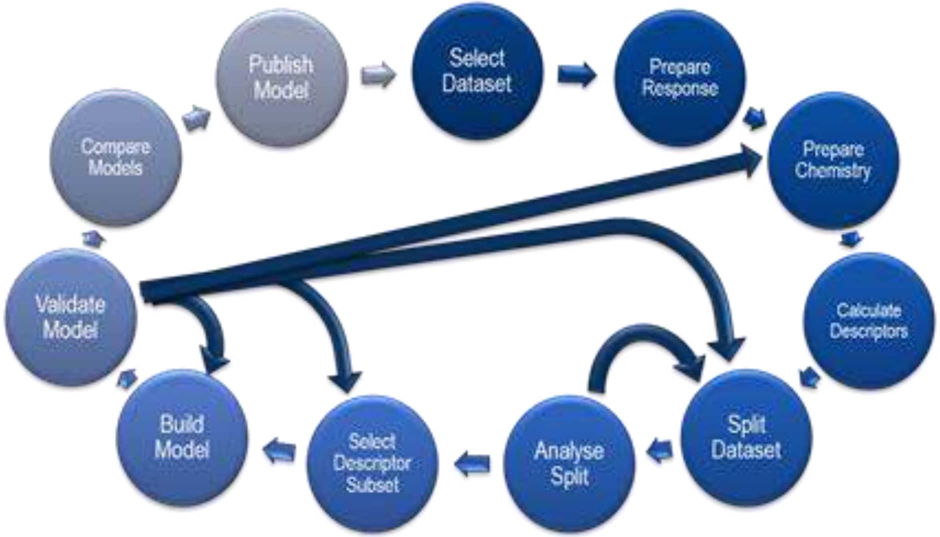
$$\text{Smooth muscle apoptosis} = 2.26(\pm 0.72) - 10.73(\pm 1.05) I_{Fe_2O_3} - 5.57(\pm 0.98) I_{\text{dextran}} - 3.53(\pm 0.54) I_{\text{surface charge}},$$

Quantitative Nanostructure-Toxicity Relationship



- Assess predictivity
- Link to mechanisms
- Refine the model

Methodology



- Tox Dataset I: Neutro Ball Cells
- Descriptors
  - Experimental setup
  - NMs Phys-Chem properties
  - Computed
- Set of 44 NMs
- ML models (GFA, NNs, etc.)

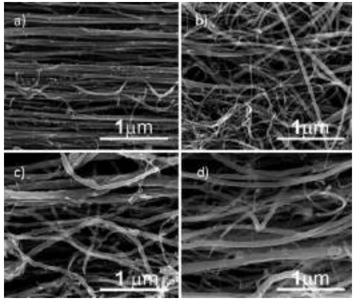
- Tox Dataset II: Inflammation
- Descriptors
  - NMs Phys-Chem properties
  - Computed Eads
- Set of 35 NMs
- ML models (GFA, NNs, etc.)

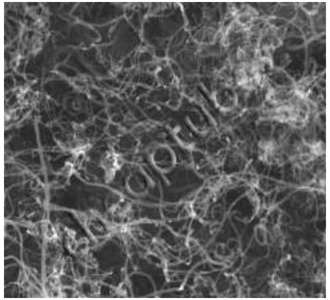
Bill of Materials…

Nanomaterials	Registration	NM ID CODE		Size data (nm)	Remarks
Mitsui-7; MWCNT; lot # 061220-24	KAJ	NRCWE-006		5 um long	powder
Cheaptubes MWCNT	KAJ	NRCWE-007		15 nm long	powder
TiO2-Rutile, 99% purity, stock#5485MR	ATS	NRCWE-024		50 nm	powder
TiO2 Pure, Pro.no:Tip-02-30	ATS	NRCWE-025		80 nm	powder
Pristine	PJA	NRCWE-040		8-15 nm	powder
Functionalized -OH	PJA	NRCWE-041		8-15 nm	powder
Functionalized -COOH	PJA	NRCWE-042		8-15 nm	powder
Pristine	PJA	NRCWE-043		>50 nm	powder
Functionalized -OH	PJA	NRCWE-044		>50 nm	powder
Functionalized -COOH	PJA	NRCWE-045		>50 nm	powder
Pristine	PJA	NRCWE-046		13-18 nm	powder
Functionalized -OH	PJA	NRCWE-047		13-18 nm	powder
Functionalized -COOH	PJA	NRCWE-048		13-18 nm	powder
Functionalized -NH2	PJA	NRCWE-049		13-18 nm	powder
Pristine >95 %, ultrapure	KBK	NRCWE-051		OD: 1-2nm; length: 5-30µm	powder
Pristine >90%, pure	KBK	NRCWE-052		OD: 1-2nm; length: 5-30µm	powder
Functionalized -OH, 3.96wt%, pure:>90%	KBK	NRCWE-053		OD: 1-2nm; length: 5-30µm	powder
Functionalized -COOH,2.73wt%, pure:>90%	KBK	NRCWE-054		OD: 1-2nm; length: 5-30µm	powder
Pristine >90%, pure; short	KBK	NRCWE-055		OD: 1-2nm; length: 1-3µm	powder
Functionalized -OH, 3.96wt%, pure:>90%, short	KBK	NRCWE-056		OD: 1-2nm; length: 1-3µm	powder
Functionalized -COOH,2.73wt%, pure:>90%, short	KBK	NRCWE-057		OD: 1-2nm; length: 1-3µm	powder
Graphene Oxide (4mg/mL, water dispension 250 ml) GO-4-250	UBV	NRCWE-058		4mg/mL	Liquid
reduced Graphene Oxide GO-4-250	UBV	NRCWE-059		50-100 micr	powder
reduced Graphene Oxide GO-4-250	UBV	NRCWE-060		125-500 micr	powder
Functionalized -NH2, 0.45wt%, purity:>95%	KBK	NRCWE-061		OD: 8-15nm; length: >50µm	powder
Pristine purity: >95 %, ultrapure	KBK	NRCWE-062		OD: <8nm; length: 10-30µm	powder
Functionalized -OH, 5.58wt%, pure:>95%	KBK	NRCWE-063		OD: <8nm; length: 10-30µm	powder
Functionalized -COOH, 3.86wt%, pure:>95%	KBK	NRCWE-064		OD: <8nm; length: 10-30µm	powder
Anatase		NM-100		Ok	m?lk (hvid)
Anatase		NM-101		Ok	M?lk(hvid)
Anatase		NM-102			M?lk(hvid)
hydrophobic rutile		NM-103			Sediments
TiO2	HWA	NRCWE-001			powder
TiO2, Rutile NaBond 80 nm (ENPRA)	HWA	NRCWE-004		80 nm	powder
Silica		NM-200			Transparent
Silica		NM-203			Transparent
MWCNT		NM-400			
MWCNT		NM-401			
difficult to disperse - SWCNT		NM-410			
difficult to disperse - SWCNT		NM-411			
TiO2, anatase, BET specific surface area 152 m^2/g		anatase		length ≅ 200-500 nm, diameter ≅ 8 nm	powder

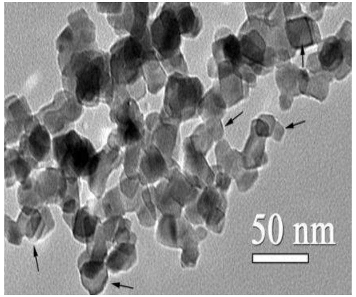
Real materials…

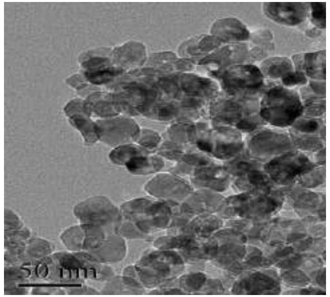
- CNTs



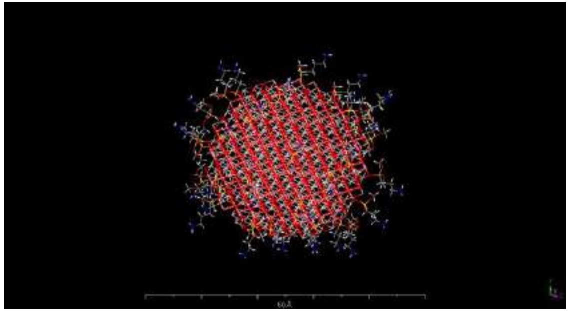


- NPs - TiO_2 – Fe_2O_3

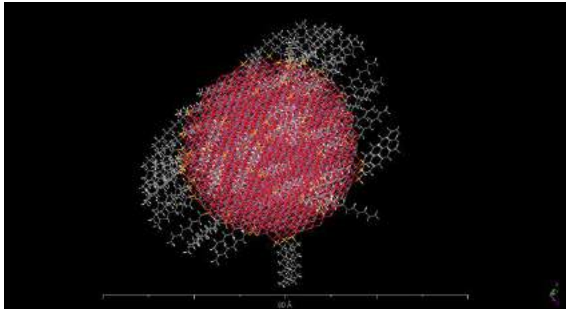




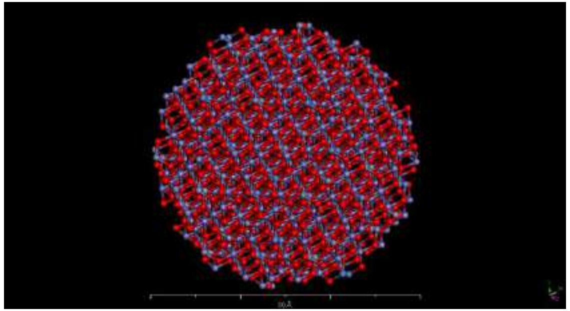
Building molecular models



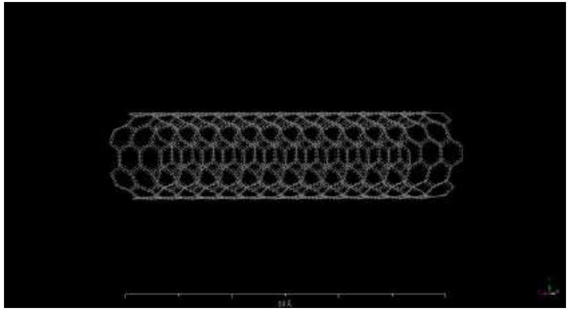
L-181



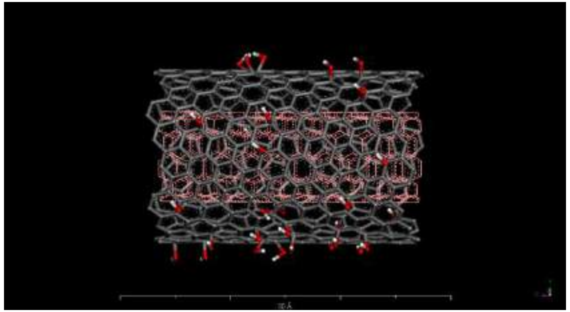
NM-111



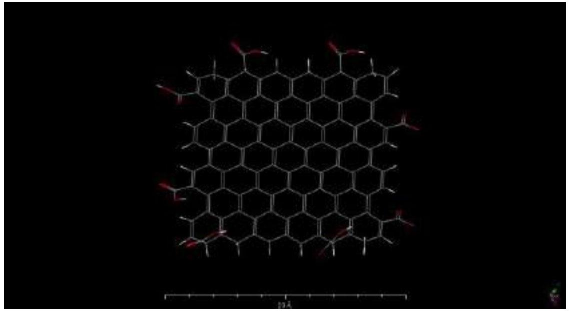
NRCWE-020



NM-411

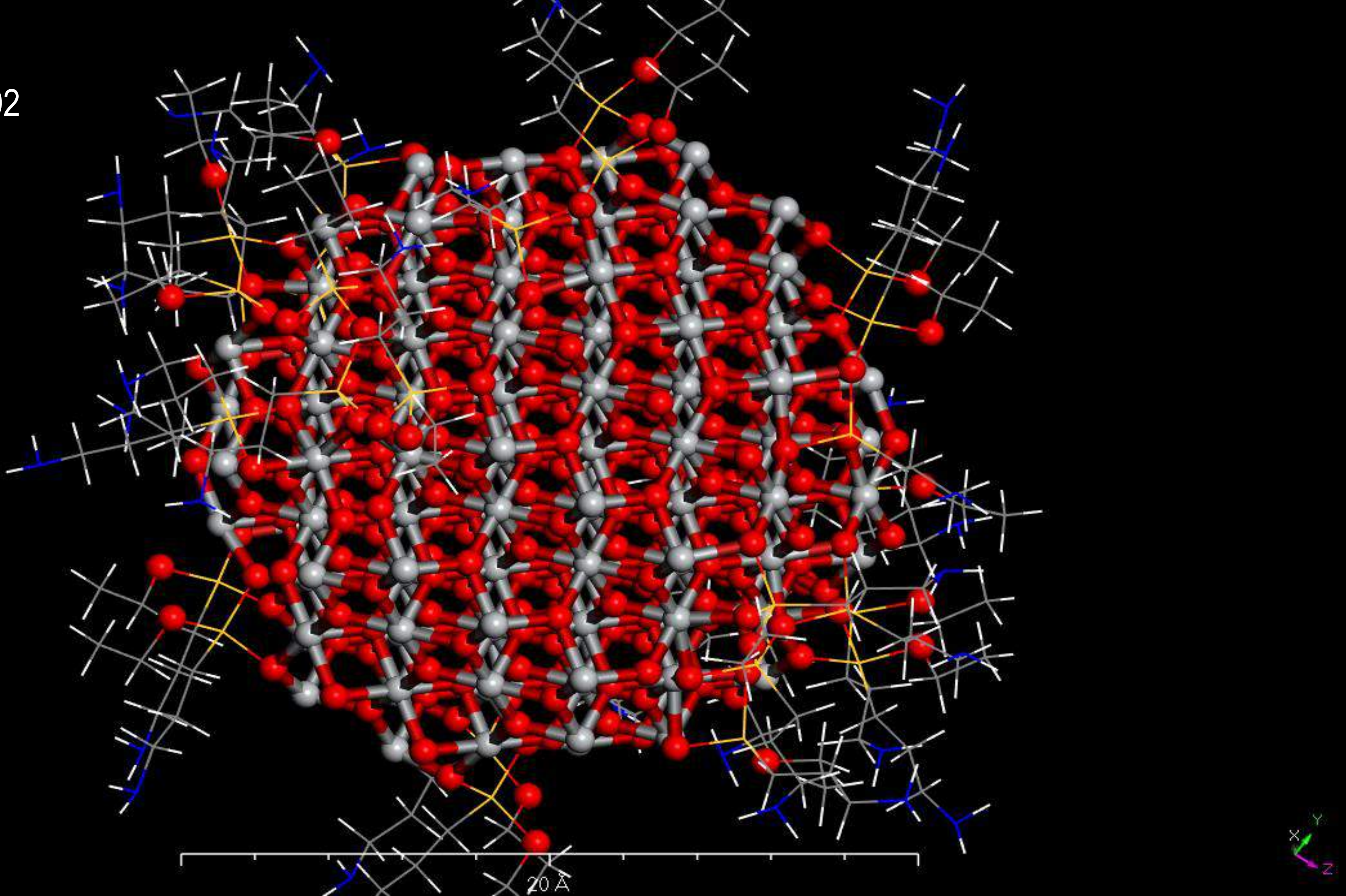


NRCWE-047



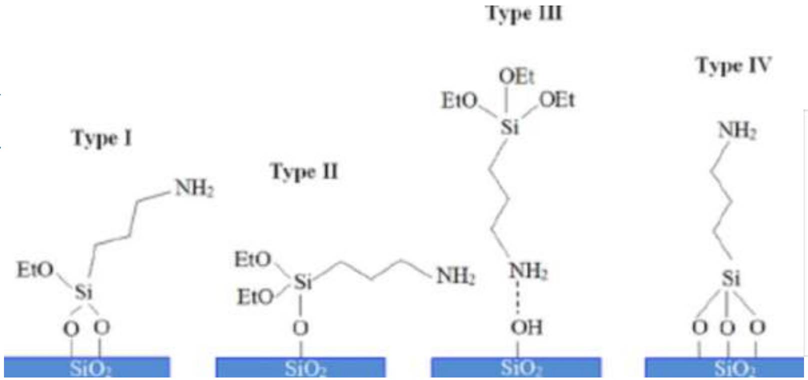
NRCWE-058

NRCWE-002

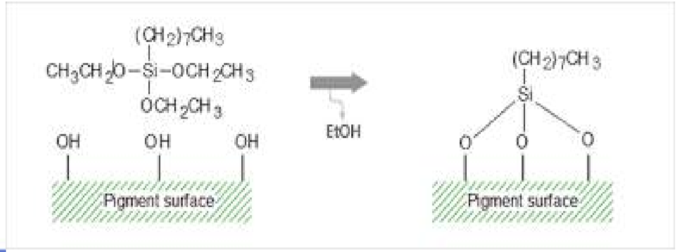


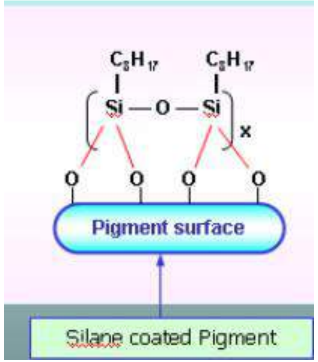
NPs				Exp. Diameter (nm)	Model size	Model Radius (nm)	Surface Chemistry		Model Size Eq.	Model Final Radius (nm)
	L181-3	TiO2	Kemira Pigments	20	10	1.5	APS	Sphere	0-10	1
	NM-110	ZnO	JRC	158	80	3	None	Sphere	11-20	1.5
	NM-111	ZnO	JRC	152	80	3* - 2	TriEthoxy CaprylylSilane	Sphere	21-30	2
	NRCWE-001	TiO2	NanoAmor	10	10	1	None	Sphere	31-40	2.5
	NRCWE-002	TiO2	NanoAmor	10	10	1	APTES (I, IV)	Sphere	41+	3
	NRCWE-018	Fe2O3	NanoAmor	40	20	2.5	None	Sphere		
	NRCWE-019	Fe2O3	NanoAmor	95	50	3	None	Cylinder		
	NRCWE-020	NiZnFe4O8	NanoAmor	20	10	1.5	None	Sphere		
	NRCWE-021	ZnFe2O4	NanoAmor	22.5	10	2	None	Sphere		
	NRCWE-022	NiFe2O4	NanoAmor	25	10	2	None	Sphere		
	NRCWE-025	TiO2	NaBond Tech.	38	20	2.5	None	Sphere		
	NRCWE-030	TiO2	NanoAmor	12	10	1.5	None	Sphere		
	PU-nTox-03	TiO2	Janez	10	10	1	None	Cylinder		
	PU-nTox-21	TiO2	Janez	20	10	1.5	None	Cube		

APTES

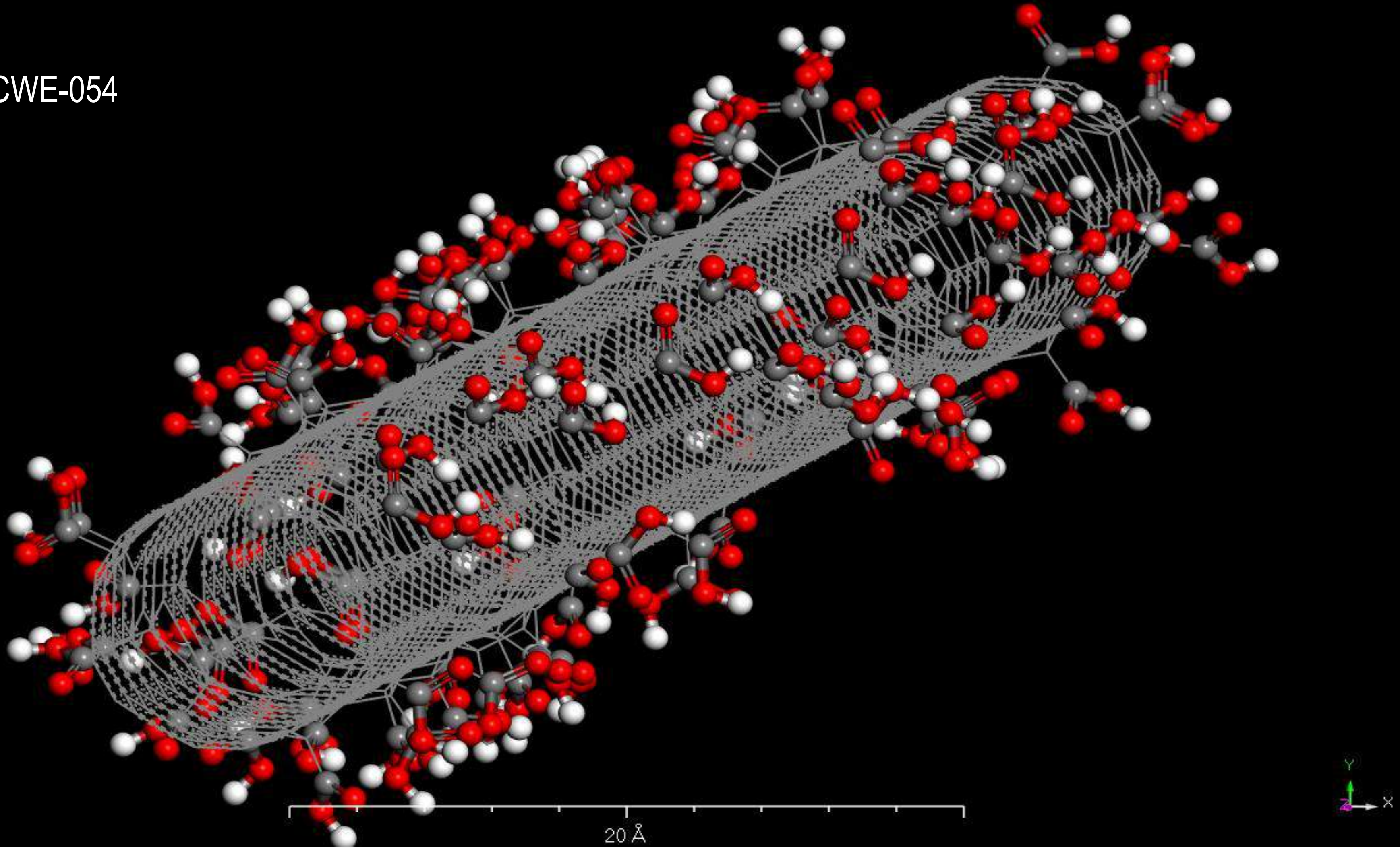


Tri Ethoxy Caprylyl Silane





NRCWE-054



TOXICITY DATA – Part I:

Y

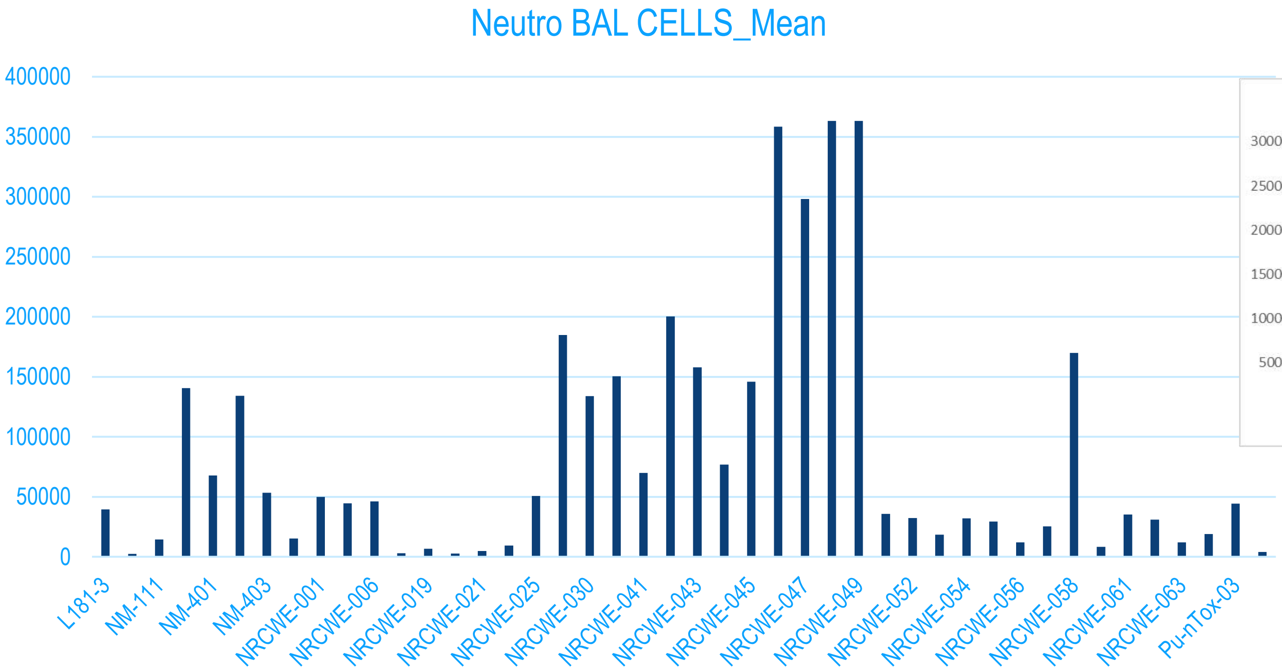
	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	NM ID CODE	Route	Dispersant	)ose (µg/anima	Sacrifice day	# Total BAL CELLS	BAL CELLSDead (not counted for total)	cro BALL CE	mpho BAL CELL	Neutro BAL CELLS	Eosino BAL CELLS	hel BAL CELLS		TL BAL COMET	%T-DNA BAL COMET	TL lung COMET	%T-DNA lung COMET	TL liver COMET	%T-DNA live
2	L181-3	Instillation	BALF	18	1	66866.66667		54496.3	2006	1671.666667	0	8692.7		25.4	9.4	16.22	9.6	12.56	3
3	L181-3	Instillation	BALF	18	1	85000		71825	0	7225	0	5950		18.1	4.57	19.63	10	17.02	3.8
4	L181-3	Instillation	BALF	18	1	64600		58786	646	969	323	3876		17	4.96	19.86	10.1	12.79	3
5	L181-3	Instillation	BALF	18	1	71400		53907	357	7140	1428	8568		27.4	8.55	21.75	10.2	19.29	7.3
6	L181-3	Instillation	BALF	18	1	117866.6667		94882.7	3536	10018.66667	0	9429.3		32.8	8.23	12.61	6.6	14.22	3.4
7	L181-3	Instillation	BALF	18	1	60066.66667		50756.3	600.666667	2102.333333	0	6607.3		26.9	9.68	18.04	9.1	13.11	3.5
8	L181-3	Instillation	BALF	54	1	94066.66667		41859.7	1881.33333	40448.66667	470.3333333	9406.7		21.5	4.87	18.51	8.6	21.53	8
9	L181-3	Instillation	BALF	54	1	164333.3333		71485	821.666667	80523.33333	821.6666667	10682		23.3	7.91	19.28	10.8	11.27	2.3
10	L181-3	Instillation	BALF	54	1	151866.6667		75933.3	759.333333	68340	0	6834		29.8	7.16	23.09	13.6	29.05	12.7
11	L181-3	Instillation	BALF	54	1	83866.66667		42772	0	32708	4612.666667	3774		22.6	6.12	17.95	8	14.42	3.7
12	L181-3	Instillation	BALF	54	1	124666.6667		53606.7	1870	62333.33333	0	6856.7		23.1	4.66	22.58	11.3	13.64	3.8
13	L181-3	Instillation	BALF	54	1	117866.6667		53040	589.333333	56576	1768	5893.3		19.5	6.16	13.74	7.5	14.35	3.3
14	L181-3	Instillation	BALF	162	1	346800		38148	3468	293046	0	12138		20.6	4.58	22.88	13.3	19.4	5.6
15	L181-3	Instillation	BALF	162	1	215333.3333		39836.7	1076.66667	157193.3333	1076.666667	16150		16.4	3.61	21.22	11.3	16.15	4.2
16	L181-3	Instillation	BALF	162	1	228933.3333		50365.3	5723.33333	164832	3434	4578.7		27.5	8.71	18.83	9.3	14.97	3.3
17	L181-3	Instillation	BALF	162	1	159800		49538	0	101473	1598	7191		25.2	7.16	19.08	9	19.49	6.5
18	L181-3	Instillation	BALF	162	1	333200		56644	0	258230	3332	14994		25.9	7.08	17.23	6.7	17.37	4.4
19	L181-3	Instillation	BALF	162	1	194933.3333		13645.3	3898.66667	164718.6667	2924	9746.7		32.4	9.43	21.06	8.8	16.97	4.5
20	L181-3	Instillation	BALF	18	3	51000		44370	510	510	255	5355		29.2	10	19.8	8.1	15.75	3.9
21	L181-3	Instillation	BALF	18	3	104266.6667		95404	0	521.3333333	1042.666667	7298.7		26.9	12.4	18.86	9	25.63	9.9
22	L181-3	Instillation	BALF	18	3	49866.66667		38397.3	997.333333	3241.333333	1496	5734.7		26.8	12.2	19.41	6.6	18.04	4.6
23	L181-3	Instillation	BALF	18	3	53266.66667		46874.7	266.333333	532.6666667	0	5593		32.5	12.6	15.18	5	14.81	3.4
24	L181-3	Instillation	BALF	18	3	66866.66667		57171	1003	334.3333333	334.3333333	8024		37.7	16.9	19.2	9.9	20.86	6.6
25	L181-3	Instillation	BALF	18	3	75933.33333		66062	1898.33333	379.6666667	0	7593.3		28.1	11.4	19.59	9.7	14.45	3.6
26	L181-3	Instillation	BALF	54	3	102000		84150	510	9180	1020	7140		31	13.6	19.78	8.3	12.82	3.6
27	L181-3	Instillation	BALF	54	3	58933.33333		48030.7	884	6188	1178.666667	2652		37	17.8	21.91	11.8	19.03	5.7
28	L181-3	Instillation	BALF	54	3	61200		51714	0	5508	306	3672		44	25.4	24.42	12.9	17.47	4.3
29	L181-3	Instillation	BALF	54	3	82733.33333		61636.3	1654.66667	9928	2068.333333	7446		43	27	22.88	10.3	16.08	3.8
30	L181-3	Instillation	BALF	54	3	45333.33333		34453.3	680	3626.666667	0	6573.3		20.5	4.91	22.88	9.9	14.55	2.5
31	L181-3	Instillation	BALF	54	3	80466.66667		63568.7	804.666667	4425.666667	804.6666667	10863		43.6	19.7	17.59	8.1	16.75	3.3
32	L181-3	Instillation	BALF	162	3	213066.6667		109729	3196	76704	10653.33333	12784		39	15.8	19.78	10	20.02	9.7
33	L181-3	Instillation	BALF	162	3	205133.3333		102567	0	91284.33333	1025.666667	10257		23.2	11.2	16.75	6.1	14.67	4
34	L181-3	Instillation	BALF	162	3	206266.6667		56723.3	10313.3333	87663.33333	41253.33333	10313		31.1	14.6	15.16	5.3	15.58	4.5
35	L181-3	Instillation	BALF	162	3	201733.3333		69598	10086.6667	113979.3333	3026	5043.3		36.9	16.2	21.09	10.3	14.6	2.9
36	L181-3	Instillation	BALF	162	3	128066.6667		76840	1921	42262	0	7043.7		40.3	15.2	19.89	10.9	13.56	2.3
37	L181-3	Instillation	BALF	162	3	157533.3333		89794	2363	56712	2363	6301.3		29.5	12	19.31	7.8	14.4	2.2
38	L181-3	Instillation	BALF	18	28	96333.33333		78030	3853.33333	2408.333333	2890	9151.7		42.8	16.1	19.7	12.3	14.61	3.1
39	L181-3	Instillation	BALF	18	28	62333.33333		57035	1870	935	0	2493.3		27.2	7.14	19.49	9.5	17.34	5.6
40	L181-3	Instillation	BALF	18	28	107666.6667		96900	1076.66667	1076.666667	0	8613.3		26.6	6.82	17.42	8.5	12.14	2.8

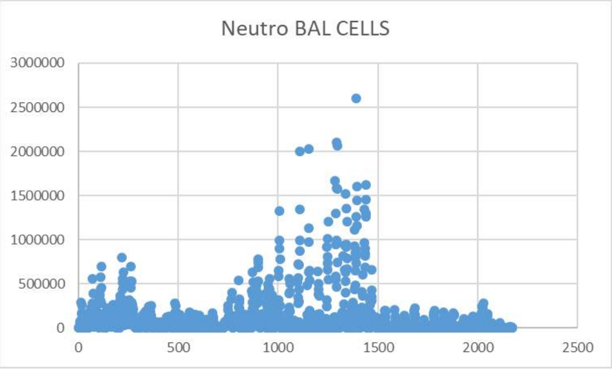
After Curation -> 2174 Data point

Experimental toxicity measurement:

NATIONAL RESEARCH CENTRE FOR THE WORKING ENVIRONMENT

HelmholtzZentrum münchen
German Research Center for Environmental Health





Mean ‘Tox’ values of each NMs

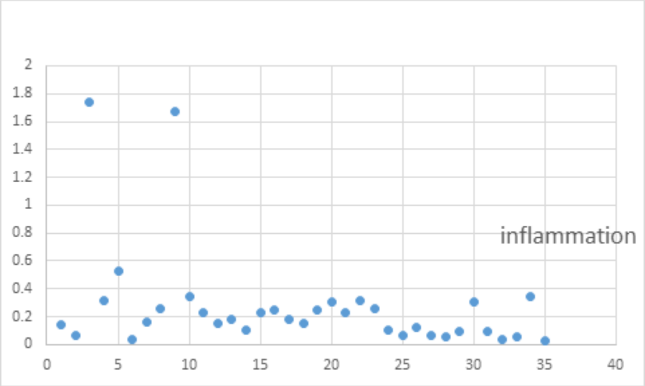
BIOVIA DASSAULT SYSTEMES

TOXICITY DATA II - Inflammation

Y

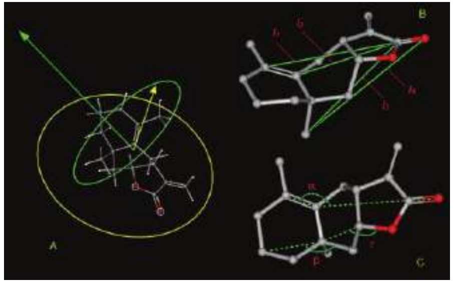
1	NAME	inflammation	Organic	ENM type	Subtype	Modification	Shape	R1-length	R2-diameter	Aspect ratio	Mean curvature	Physical surface area	BET Surface area	ROS Surface
2	L181-3	0.141	0	TiO2	Rutile	SiAl2rPolyalcohol	Spherical	38.4	20.6	1.86	0.0971	1333.166258	107.7	8
3	NM-400	0.062	1	CNT	MWCNT	Pristine	Tube	847	11	77.00	0.0909	29460.28511	254	9
4	NM-401	1.744	1	CNT	MWCNT	Pristine	Tube	4048	67	60.42	0.0149	859101.4978	18	3
5	NM-402	0.313	1	CNT	MWCNT	Pristine	Tube	1372	11	124.73	0.0909	47602.98268	226	1
6	NM-403	0.528	1	CNT	MWCNT	Pristine	Tube	443	12	36.92	0.0833	16926.90122	135	3
7	NM-411	0.037	1	CNT	SWCNT	Pristine	Tube	1000	2	500.00	0.5000	6289.468492	861	4
8	NRCWE-001	0.165	0	TiO2	Rutile	Pristine	Spherical	10	10	1.00	0.2000	314.1592654	99	2
9	NRCWE-002	0.26	0	TiO2	Rutile	Positive	Spherical	10	10	1.00	0.2000	314.1592654	84.3	4
10	NRCWE-006	1.676	1	CNT	MWCNT	Pristine	Tube	573	74	7.74	0.0135	141811.4924	26	1
11	NRCWE-025	0.345	0	TiO2	Rutile	Pristine	Spherical	26	26	1.00	0.0769	2123.716634	28.2	3
12	NRCWE-026	0.233	1	CNT	MWCNT	Pristine	Tube	846	11	76.91	0.0909	29425.72759	254	4
13	NRCWE-030	0.154	0	TiO2	Rutile	Pristine	Spherical	12.1	12.1	1.00	0.1653	459.9605804	139.1	
14	NRCWE-040	0.177	1	CNT	MWCNT	Pristine	Tube	519	22	23.48	0.0455	36800.94621	150	8
15	NRCWE-041	0.1	1	CNT	MWCNT	OH	Tube	1005	27	37.36	0.0370	86068.03052	152	1
16	NRCWE-042	0.226	1	CNT	MWCNT	COOH	Tube	723	30	23.94	0.0333	70028.04804	141	1
17	NRCWE-043	0.25	1	CNT	MWCNT	Pristine	Tube	771	27	28.84	0.0370	65866.87175	82	8
18	NRCWE-044	0.182	1	CNT	MWCNT	OH	Tube	1330	33	40.67	0.0303	138310.6429	74	9
19	NRCWE-045	0.157	1	CNT	MWCNT	COOH	Tube	1553	30	51.42	0.0333	148775.2095	119	1
20	NRCWE-046	0.252	1	CNT	MWCNT	Pristine	Tube	717	29	24.64	0.0345	66878.55428	223	2
21	NRCWE-047	0.308	1	CNT	MWCNT	OH	Tube	533	23	23.58	0.0435	38645.29672	216	2
22	NRCWE-048	0.229	1	CNT	MWCNT	COOH	Tube	1604	18	89.61	0.0556	90703.45048	185	8
23	NRCWE-049	0.315	1	CNT	MWCNT	NH2	Tube	731	15	49.06	0.0667	34566.64552	199	1
24	NRCWE-051	0.254	1	CNT	SWCNT	Pristine	Tube	17500	21	829.38	0.0476	1160732.422	442.6	1
25	NRCWE-052	0.101	1	CNT	SWCNT	Pristine	Tube	17500	19	925.93	0.0526	1039642.874	405.7	1
26	NRCWE-053	0.069	1	CNT	SWCNT	OH	Tube	17500	24	729.17	0.0417	1320373.693	367.8	3
27	NRCWE-054	0.126	1	CNT	SWCNT	COOH	Tube	17500	17	1005.75	0.0588	957090.5373	370.8	1
28	NRCWE-055	0.067	1	CNT	SWCNT	Pristine	Tube	2000	14	145.99	0.0714	86374.46147	453.1	5
29	NRCWE-056	0.058	1	CNT	SWCNT	OH	Tube	2000	13	156.25	0.0769	80682.1312	373.4	4
30	NRCWE-057	0.09	1	CNT	SWCNT	COOH	Tube	2000	7	298.51	0.1429	42167.85461	281.6	7
31	NRCWE-061	0.303	1	CNT	MWCNT	NH2	Tube	731	16.42196296	44.51	0.0609	38136.72112	170.4	1
32	NRCWE-062	0.099	1	CNT	MWCNT	Pristine	Tube	468	8.818419643	53.07	0.1134	13087.56918	443.2	2
33	NRCWE-063	0.041	1	CNT	MWCNT	OH_more	Tube	345	14.17828767	24.33	0.0705	15682.89703	426.4	1
34	NRCWE-064	0.053	1	CNT	MWCNT	COOH_more	Tube	214	7.464317757	28.67	0.1340	5105.785558	445.2	5
35	PU-nTox-03	0.346	0	TiO2	anatase	Pristine	Tube	350	10	35.00	0.1000	11152.65392	152	6
36	PU-nTox-21	0.032	0	TiO2	anatase	Pristine	Cube	20	20	1.00	0.0000	2400	86	1

Inflammation efficacy in units cm^2/g for the lung



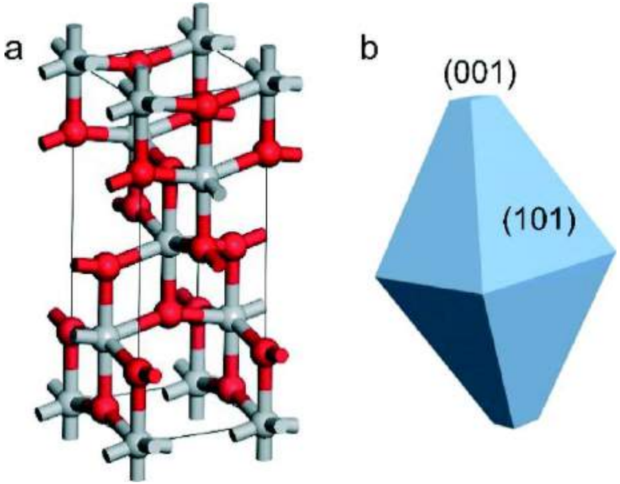
Using Log (Inflammation) as Y

NMs DESCRIPTORS – Classes and Types



Challenge = NMs are ‘substances’ not ‘small molecules’ (typically used in QSAR studies)

- Experimental – Characterisation
  - Shape
  - …
- Computed by Simulation (**WP4**)
  - Redox Potentials
  - Adsorption Energies on
  - Band Gap
  - …
- Common descriptors (standard?)
  - Log P
- Descriptors derived from Exp. Protocols



Descriptors (Independent Variables)

- Two subsets of NMs
  - Subset I: 44 NMs used in the TOX dataset (NRCWE)
  - Subset II: 35 NMs used in the Inflammation dataset (UCD)
- Experimental Descriptors for the NMs used are the same for both datasets
- Computed Descriptors
  - For Subset II – add computed Adsorption energies on proteins, AAs and Lipids (DPPC)

Descriptors – Experimental (NanoReg 2 DB)

- Physico-chemical properties (before/after)
- Well defined characterization techniques are still a challenge

- ✓ Size and shape
- ✓ Size distribution
- ✓ Agglomeration state
- ✓ Porosity
- ✓ Structure-dependent elect.
- ✓ Electronic properties
- ✓ **Characterisation:**

  In what medium?

- ✓ Surface area
- ✓ Surface chemistry
- ✓ Surface charge
- ✓ Crystal structure
- ✓ Composition
- ✓ Configuration
- ✓ **Coating**
- ✓ **Aggregation state**
- ✓ Metal content

Toxicology and Applied Pharmacology 299 (2016) 95–100

Contents lists available at ScienceDirect

Toxicology and Applied Pharmacology

journal homepage: www.elsevier.com/locate/ytaap

Recent advances, and unresolved issues, in the application of computational modelling to the prediction of the biological effects of nanomaterials☆

David A. Winkler

Particuology 21 (2015) 1–19

Contents lists available at ScienceDirect

Particuology

journal homepage: www.elsevier.com/locate/partic

Invited review

(Q)SAR modelling of nanomaterial toxicity: A critical review

Ceyda Oksel[a], Cai Y. Ma[a], Jing J. Liu[a,b], Terry Wilkins[a], Xue Z. Wang[a,b,*]

[a] Institute of Particle Science and Engineering, School of Chemical and Process Engineering, University of Leeds, Leeds LS2 9JT, UK
[b] School of Chemistry and Chemical Engineering, South China University of Technology, Guangzhou 510641, China

Chapter 5
Literature Review of (Q)SAR Modelling of Nanomaterial Toxicity

Ceyda Oksel, Cai Y. Ma, Jing J. Liu, Terry Wilkins, and Xue Z. Wang

The pro-inflammatory effects of low-toxicity low-solubility particles, nanoparticles and fine particles, on epithelial cells in vitro: the role of surface area

Claire Monteiller, Lang Tran, William MacNee, Steve Faux, Alan Jones, Brian Miller, and Ken Donaldson

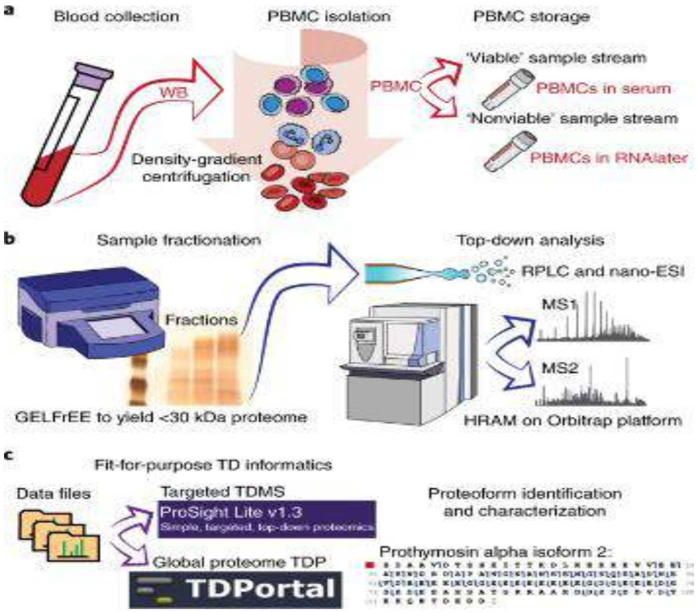
BIOVIA

DASSAULT SYSTEMES

Descriptors – Experimental - continued

Experimental Protocols:

- DOSAGE, Environment, Spectral information (NMR, UV,…)
- Sacrifice day, Route, Dispersant, etc.



ELSEVIER

Food and Chemical Toxicology

Volume 112, February 2018, Pages 478-494

Perspectives from the NanoSafety Modelling Cluster on the validation criteria for (Q)SAR models used in nanotechnology

Tomasz Puzyn [a,1], Nina Jeliazkova [b], Haralambos Sarimveis [c], Richard L. Marchese Robinson [d,1,2], Vladimir Lobaskin [e], Robert Rallo [f], Andrea-N. Richarz [d], Agnieszka Gajewicz [a], Manthos G. Papadopulos [g], Janna Hastings [h,3], Mark T.D. Cronin [d], Emilio Benfenati [i], Alberto Fernández [j]

Show more

BIOVIA

DASSAULT SYSTEMES

Descriptors (2) – Computed

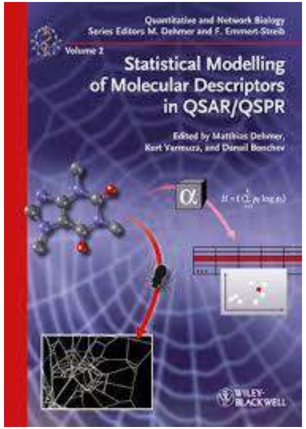
See WP4

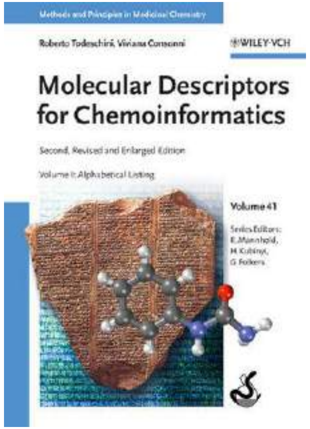
- Simulations:
  - Hydration energies
  - Bio-nano binding free energies (Eads)
- Others:
  - Surface area
  - Volume, Shape, Surface area…
  - Dipole Moment
  - Atom Count

Use of Metal Oxide Nanoparticle Band Gap to Develop a Predictive Paradigm for Oxidative Stress and Acute Pulmonary Inflammation

Haiyuan Zhang[1,*], **Zhaoxia Ji**[1,*], **Tian Xia**[2], **Huan Meng**[2], **Cecile Low-Kam**[3], **Rong Liu**[4], **Suman Pokhrel**[5], **Sijie Lin**[1], **Xiang Wang**[1], **Yu-Pei Liao**[2], **Meiying Wang**[2], **Linjiang Li**[1], **Robert Rallo**[6], **Robert Damoiseaux**[1,7], **Donatello Telesca**[3], **Lutz Mädler**[5], **Yoram Cohen**[4], **Jeffrey I. Zink**[8], and **Andre E. Nel**[1,2,$]

[1]California NanoSystems Institute, University of California, Los Angeles, California

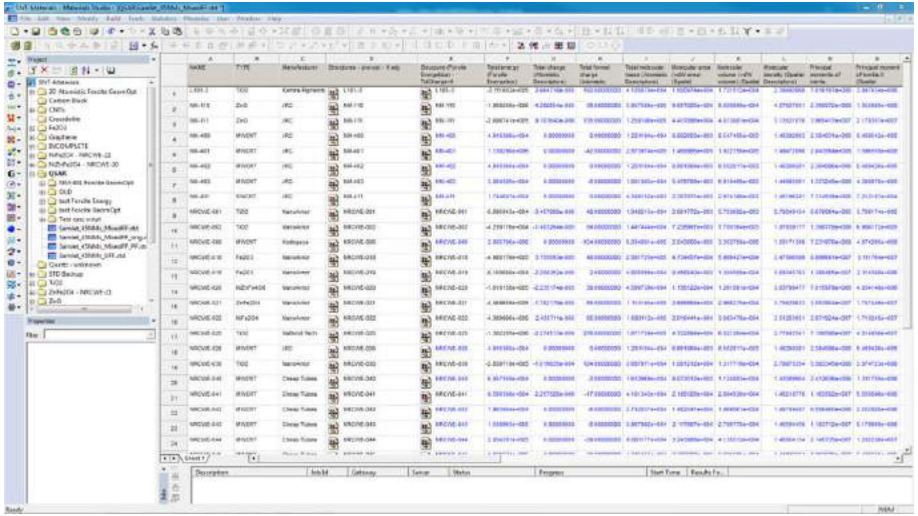




BIOVIA Materials Studio

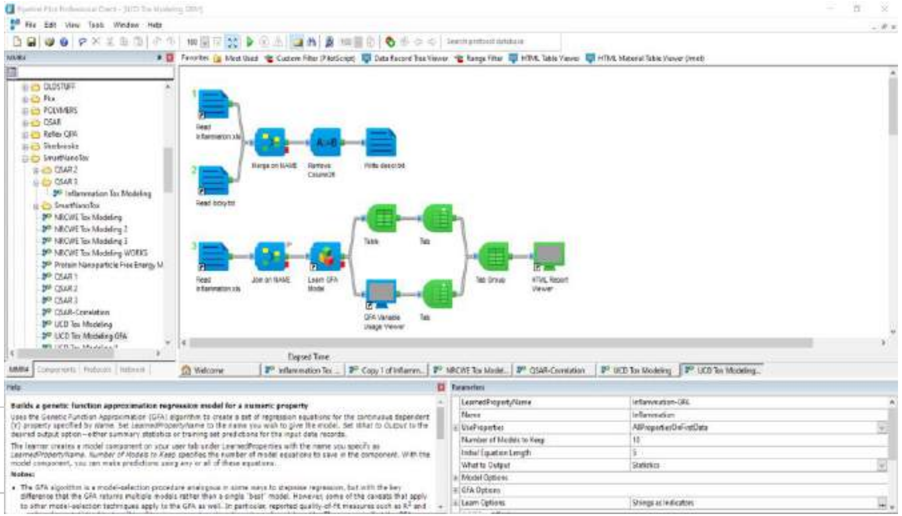
Solve key materials and chemical research problems with an integrated, multi-scale modeling environment that delivers a complete range of simulation methods.

BIOVIA
MODEL THE BIOSPHERE

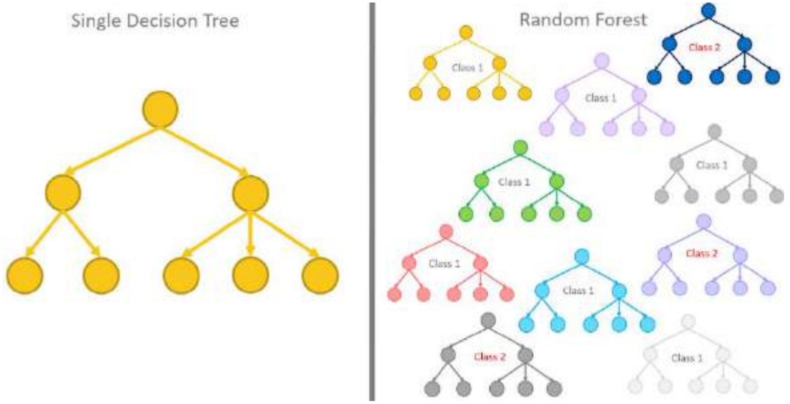


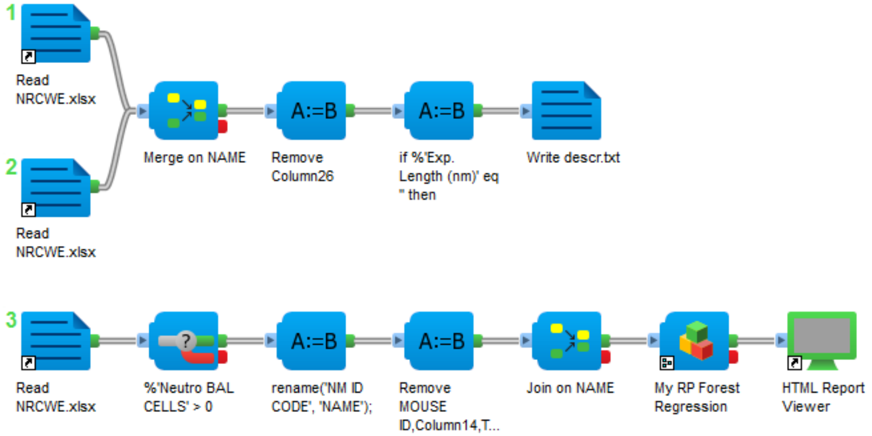
BIOVIA Pipeline Pilot

Data Science Tools for Scientists



Tox Dataset 1 - RP Forest Models





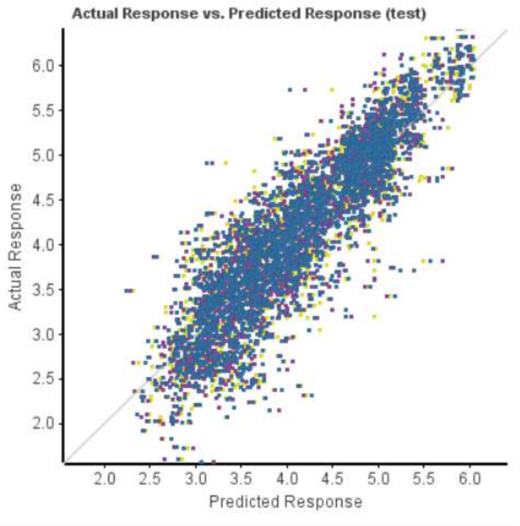
Using RP Forest model
Y = Neutro BALL Cell (%, Log)

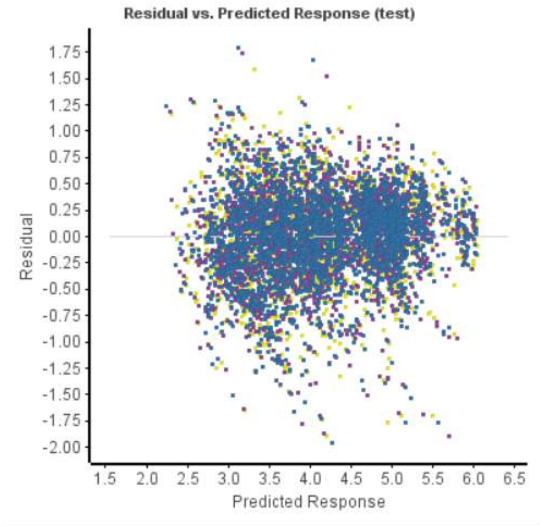
[Learn RP Forest Model – Split Method = Gini – Max. tree depth 100 – Cross-Validation = Random]

Tox Dataset 1 QSAR model - Details

Summary | Plots | Importance

Summary Statistics							
R-squared (training)	RMS Error (training)	Iteration	Q-squared (test)	RMS Error (test)	RMS Error (null model)	Overall Q-squared (test)	Overall RMS Error (test)
0.85071	0.35871	1 2 3	0.79663 0.80036 0.80013	0.41592 0.41209 0.41245	0.91993 0.91950 0.91937	0.79904	0.41349

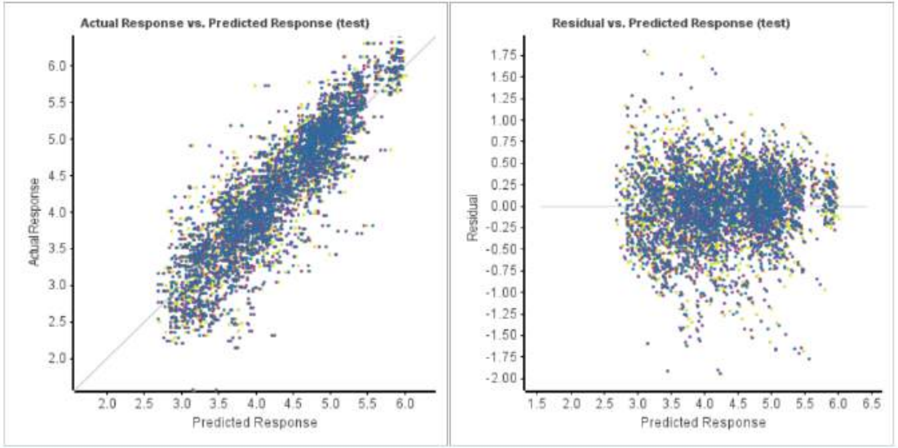




Importance - Iteration 3 (test)				
Variable Name	Variable Importance (RMS Prediction Change)	Variable Importance (RMSE Change)	Variable Importance (R2 Change)	Variable Correlation with Response
Sacrifice day	0.6713	0.386	0.47621	-0.22767
Dose (µg/animal)	0.5196	0.2787	0.33694	0.28246
TYPE	0.284	0.09245	0.09216	NA
C_Count	0.1748	0.03929	0.03094	0.07407
Dispersant	0.1692	0.03425	0.03136	NA
Shadow area fraction: ZX plane (Spatial Descriptors)	0.09329	0.01023	0.00592	0.26041
Route	0.0881	0.009703	0.00825	NA
Shadow area fraction: XY plane (Spatial Descriptors)	0.08716	0.01068	0.00596	0.27274
BET	0.08565	0.009199	0.00614	0.02512
Manufacturer	0.07306	0.006721	0.00483	NA
Exp. Diameter (nm)	0.06891	0.006468	0.00427	-0.17562
Ti_Count	0.066	0.007099	0.00600	0.00293
H_Count	0.05801	0.00474	0.00125	0.06471
Fe_Count	0.05483	0.004287	0.00284	-0.26133
O_Count	0.05212	0.00478	0.00280	-0.24779
Shadow ratio (Spatial Descriptors)	0.04909	0.00216	-0.00011	-0.05619
Exp. Length (nm)	0.04846	0.003424	0.00104	-0.02849
Shadow length: LX (Spatial Descriptors)	0.03993	0.002094	0.00067	-0.06569
Dipole moment Y (Spatial Descriptors)	0.03865	0.001658	0.00080	-0.01214
Radius of gyration (Spatial Descriptors)	0.03862	0.001063	-0.00019	-0.05921
Shadow area fraction: YZ plane (Spatial Descriptors)	0.03674	0.00282	0.00181	-0.00136
Shadow area: YZ	0.03646	0.001896	0.00072	-0.03070

Tox Dataset 1 QSAR model 2 –adding Eads Descriptors

Summary Statistics							
R-squared (training)	RMS Error (training)	Iteration	Q-squared (test)	RMS Error (test)	RMS Error (null model)	Overall Q-squared (test)	Overall RMS Error (test)
0.84274	0.34799	1	0.79155	0.39737	0.86846	0.79035	0.39849
		2	0.78904	0.39984	0.86773		
		3	0.79048	0.39825	0.86783		



Importance - Iteration 3 (test)				
Variable Name	Variable Importance (RMS Prediction Change)	Variable Importance (RMSE Change)	Variable Importance (R2 Change)	Variable Correlation with Response
Sacrifice day	0.7143	0.4371	0.56295	-0.31994
Dose (µg/animal)	0.522	0.2848	0.36485	0.29108
Dispersant	0.1765	0.04215	0.04490	NA
Route	0.09067	0.01138	0.01070	NA
BET	0.08464	0.007371	0.00648	-0.14162
Eads(SER)	0.04556	0.00385	0.00177	-0.26558
Eads(DPPC)	0.04481	0.00261	0.00087	-0.19775
Eads(1YCK)	0.0428	0.003508	0.00162	-0.21929
Eads(5WY9)	0.04242	0.003154	0.00143	-0.21908
Eads(5NRF)	0.04228	0.00325	0.00165	-0.23806
Eads(LEU)	0.04098	0.002922	0.00094	-0.25771
Eads(6ENP)	0.0401	0.00345	0.00178	-0.22640
Eads(CYS)	0.03918	0.002022	0.00024	-0.27420
Eads(1NDX)	0.03688	0.002403	0.00097	-0.22541
Eads(4XAT)	0.03608	0.001857	0.00017	-0.21033
Eads(4GLP)	0.03495	0.00181	0.00017	-0.20396
Eads(2PKT)	0.03472	0.001489	-0.00025	-0.20304
Eads(spd)	0.03339	0.001195	-0.00013	-0.18147
Eads(4MTH)	0.03287	0.001307	-0.00019	-0.21758
H_Count	0.03216	0.001487	0.00107	-0.00019
Eads(THR)	0.03127	0.001303	-0.00023	-0.26404
Eads(ASP)	0.03017	0.0007721	0.00029	0.12478
Eads(5E13)	0.02952	0.0006674	-0.00069	-0.22747
Eads(ETA)	0.02727	0.0002576	-0.00016	0.17940
Eads(5LG8)	0.02647	0.0007715	-0.00005	-0.23246
Manufacturer	0.02484	0.00162	0.00136	NA

Variables Correlation Matrix

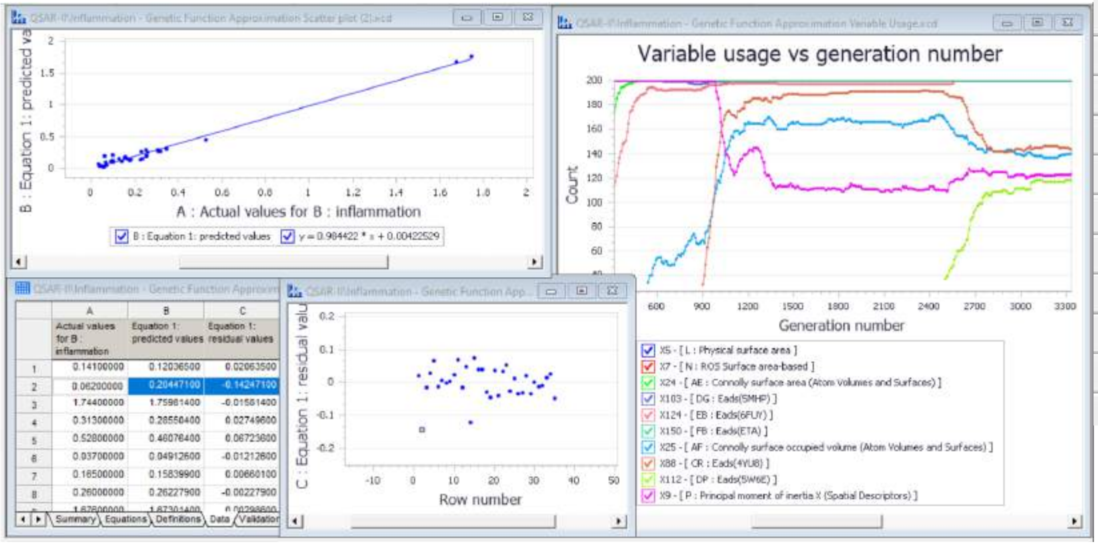
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN
	I: R1-I	J: R2-d	K: Aspe ratio	L: Mear curv	M: Phys surf	N: BET Surf	O: ROS Surf	P: Princ mom	Q: Princ mom	R: Princ mom	S: Princ mom	T: Radi of	U: Ellips volur	V: Shac area	W: Shac area	X: Shac area	Y: Shac area	Z: Shac area	AA: Shac area	AB: Shac lengt	AC: Shac lengt	AD: Shac lengt	AE: Shac ratio	AF: Conr surf	AG: Conr surf	AH: Solv surf	AI: Solv surf	AJ: Dipol mom	AK: Dipol mom	AL: Dipol mom	AM: Dipol mom	AO: H_C	AP: C_C	AQ: N_C	AR: O_C	AS: Na_C	AT: Si_C	AU: Ti_C	AV: Fe_C
O : ROS	8800	1050	4200	2700	6000	7910	1	3040	2900	2250	4070	9830	1400	9610	2100	4430	6500	7700	3200	7960	6900	1700	0160	9540	5200	9140	5400	4230	1610	6900	2900	1100	9590	2298	4600	0	3400	1100	0
P : Principal	9420	9170	8140	7400	8010	6550	3040	1	0700	8800	8800	2600	8700	8600	9620	7300	0100	0470	5200	7800	2220	1440	2700	5200	0300	5600	0500	5600	1100	3000	0500	0400	6200	5600	3020	0	7780	4980	0
Q : Principal	7030	8900	5800	0900	7740	3800	2900	0700	1	7670	6880	6340	6800	9000	5900	0200	8300	7000	5700	5330	2600	7000	5200	2600	2000	7900	5300	5180	6470	1560	4770	5300	8200	7240	-004	0	8710	2230	0
R : Principal	6410	7640	5360	6750	6290	2750	2250	8800	7670	1	1	5300	3200	5200	5480	1700	9980	9340	9680	4200	5650	7710	4100	1000	9700	9600	8300	6600	2700	2700	4000	2700	5700	2400	7820	0	5880	9140	0
S : Principal	8640	0170	6280	9280	8220	7010	4070	8800	6880	1	1	3800	2700	4800	7790	1200	0640	2950	3140	2400	9630	1860	1200	1500	9600	9400	8300	7700	2000	1200	5800	2700	5300	2700	6800	0	5680	8440	0
T : Radius of	0900	3175	7800	2800	9700	9100	9830	2600	6340	5300	3800	1	5300	7000	6250	6400	8830	6920	2140	6500	3240	7460	4900	9800	0200	9200	8100	1100	8100	2000	6700	2500	0300	1900	5170	0	8400	2700	0
U : Ellipsoidal	3000	0200	1600	8400	3930	8900	1400	8700	6800	3200	2700	5300	1	1900	9500	2400	3200	6900	5600	4500	7000	6300	5700	2500	0900	8100	9900	7800	5000	2100	3800	3500	0900	7200	6390	0	0560	7980	0
V : Shadow	3680	1500	5070	7500	9200	5420	9610	8600	9000	5200	4800	7000	1900	1	2300	5800	4700	3800	2600	7100	6900	5800	9200	4800	2900	6200	1700	3600	5800	9800	8100	0700	9500	2100	7159	0	0500	5330	0
W : Shadow	1600	8300	1700	1400	9910	2300	2100	9620	5900	5480	7790	6250	9500	2300	1	1900	6200	7900	2100	7900	1100	9100	0800	7900	6700	5300	8500	1440	2160	9605	1050	2700	2400	1370	5900	0	6050	8100	0
X : Shadow	6910	6800	0080	0100	2100	7930	4430	7300	0200	1700	1200	6400	2400	5800	1900	1	4100	3300	6700	6500	1000	2400	6700	0500	9200	3100	1300	7600	3700	2100	0900	9300	9000	6500	2640	0	1700	2490	0
Y : Shadow	6500	4900	6400	3630	6200	0800	6500	0100	8300	9980	0640	8830	3200	4700	6200	4100	1	0000	8900	8490	0500	4350	4850	0800	2600	5600	2600	0360	8050	8820	2600	7300	5500	9900	9600	0	9500	6600	0
Z : Shadow	2700	4100	4600	7700	6700	4600	7700	0470	7000	9340	2950	6920	6900	3800	7900	3300	0000	1	6260	6480	5400	6400	4000	3800	0900	6800	9400	2700	0320	1530	3500	5100	6900	6570	1100	0	9800	3200	0
AA : Shadow	1100	4100	4200	9220	2700	0400	3200	5200	5700	9680	3140	2140	5600	2600	2100	6700	8900	6260	1	4710	0280	5200	5350	5600	6700	7800	8200	2330	0800	5290	5500	9700	0000	9900	2700	0	2900	1500	0
AB : Shadow	2700	6580	4500	8500	8300	8300	7960	7800	5330	4200	2400	6500	4500	7100	7900	6500	8490	6480	4710	1	6700	4600	2300	7100	3500	7700	4900	2800	7300	2100	6500	6000	2800	1300	3710	0	9500	4900	0
AC : Shadow	5000	3300	0200	9500	7960	8000	6900	2220	2600	5650	9630	3240	7000	6900	1100	1000	0500	5400	0280	6700	1	9200	9100	2900	6400	5800	7000	9450	1812	1050	0010	3200	8100	7186	9600	0	0000	4100	0
AD : Shadow	0400	2700	3400	9000	6530	2200	1700	1440	7000	7710	1860	7460	8300	5800	9100	2400	4350	6400	5200	4600	9200	1	4700	3300	6900	9600	8000	8900	1897	2780	2240	4400	6200	6450	7900	0	6800	5700	0
AE : Shadow	2000	5800	6500	3550	5700	3000	0160	2700	5200	4100	1200	4900	5700	9200	0800	6700	4850	4000	5350	2300	9100	4700	1	3800	9650	7800	2000	8700	6400	0500	1400	5400	4340	5000	7300	0	9000	7000	0
AF : Connolly	0900	5800	5250	8100	6000	3400	9540	5200	2600	1000	1500	9800	2500	4800	7900	0500	0800	3800	5600	7100	2900	3300	3800	1	6700	0100	5200	3200	2700	7000	0000	9900	7000	7300	4696	0	8000	6560	0
AG : Connolly	5290	7300	7800	3200	8220	2600	5200	0300	2000	9700	9600	0200	0900	2900	6700	9200	2600	0900	6700	3500	6400	6900	9650	8700	1	7000	4500	5600	9100	4130	8840	4500	4000	9430	3090	0	4540	6480	0
AH : Solvent	6200	9900	9170	8000	9400	8300	9140	5600	7900	9600	9400	9200	8100	6200	5300	3100	5600	6800	7800	7700	5800	9600	7800	0100	7000	1	4400	0900	4600	6200	0600	2700	5500	7300	0185	0	0670	7281	0
AI : Solvent	0250	1100	1900	4500	3900	0000	5400	0500	5300	8300	6300	8100	9900	1700	8500	1300	2600	9400	8200	4900	7000	8000	2000	5200	4500	4400	1	7700	3900	1570	7910	2900	4800	6500	1400	0	1260	8740	0
AJ : Dipole	1500	9800	6800	4600	5200	1200	4230	5600	5180	6600	7700	1100	7800	3600	1440	7600	0360	2700	2330	2800	9450	8900	8700	3200	5600	0900	7700	1	2200	8270	9100	4990	8950	0200	8900	0	5090	3300	0
AK : Dipole	5340	4770	2310	4200	2530	3870	1610	1100	6470	2700	2000	8100	5000	5800	2160	3700	8050	0320	0800	7300	1812	1897	6400	2700	9100	4600	3900	2200	1	7800	2600	7100	5600	4300	1762	0	0160	2390	0
AL : Dipole	8340	5690	9870	7700	2078	3380	6900	3000	1560	2700	1200	2000	2100	9800	9605	2100	8820	1530	5290	2100	1050	2780	0500	7000	4130	6200	1570	8270	7800	1	0200	0390	4180	9500	4560	0	1630	3560	0
AM : Dipole	9870	6790	6450	9940	6490	6290	2900	0500	4770	4000	5800	6700	3800	8100	1050	0900	2600	3500	5500	6500	0010	2240	1400	0000	8840	0600	7910	9100	2600	0200	1	2500	7700	2400	6200	0	9280	3600	0
AO : H_Count	0110	8200	1500	1800	4520	6000	1100	0400	5300	2700	2700	2500	3500	0700	2700	9300	7300	5100	9700	6000	3200	4400	5400	9900	4500	2700	2900	4990	7100	0390	2500	1	3000	5900	5000	0	2430	9000	0

Highly correlated variables (>=0.7) are removed from the list

e.g. Spatial descriptors like Volume, Length, Moment Inertia or Atom counts (C&H, Ti&O)

For the Eads (Dataset II) only used 1 variable for Protein Adsorption, 2 for small mol. Ads (AA and Lipids)

GFA Model on Tox II (Inflammation)



	Equation 1	Equation 2	Equat
Friedman LOF	1.26712000	1.28411900	1.
R-squared	0.73045200	0.72683600	0.
Adjusted R-squared	0.67269200	0.66830100	0.
Cross validated R-squared	0.49424200	0.53861400	0.
Significant Regression	Yes	Yes	Yes
Significance-of-regression F-value	12.64626400	12.41707600	12.
Critical SOR F-value (95%)	2.45854700	2.45854700	2.
Replicate points	0	0	
Computed experimental error	0.0000000	0.0000000	0.
Lack-of-fit points	28	28	
Min expt. error for non-significant LOF (95%)	0.43483800	0.43774500	0.

Equation	Definitions
Y = 0.000040113 * X1 + 4.746605621 * X3 - 0.004685155 * X4 + 0.000041291 * X11 + 0.000717498 * X16 - 0.027853961 * X21 - 2.430530061	X1 : E : R1-length X3 : G : Mean curvature X4 : H : BET Surface area X11 : O : C_Count X16 : T : Ti_Count X21 : Y : Eads(DPPC)

BIOVIA DASSAULT SYSTEMES

RP Forest Regression Model on Tox Dataset II (Inflammation)

Summary | Plots | Importance

Summary Statistics							
R-squared (training)	RMS Error (training)	Iteration	Q-squared (test)	RMS Error (test)	RMS Error (null model)	Overall Q-squared (test)	Overall RMS Error (test)
0.91194	0.37856	1	0.32352	0.75984	0.97979	0.21756	0.80756
		2	0.24827	0.79127	0.95651		
		3	0.31629	0.76659	0.93642		
		4	0.11774	0.86170	0.95838		
		5	0.13042	0.85272	0.91788		

Importance - Iteration 5 (test)				
Variable Name	Variable Importance (RMS Prediction Change)	Variable Importance (RMSE Change)	Variable Importance (R2 Change)	Variable Correlation with Response
BET Surface area	0.3289	0.1202	0.12873	-0.61660
Eads(DPPC)	0.2261	0.007104	0.00938	-0.17482
R1-length	0.08208	0.008013	0.01807	-0.09415
Eads(TLR4)	0.07881	-0.003571	-0.00461	-0.26591
R2-diameter	0.07203	0.0006697	0.00546	0.64120
Eads(CYM)	0.06749	0.001253	0.00075	0.23376
ROS Surface	0.05662	-0.02305	-0.04676	0.13845
H_Count	0.04899	-0.006933	-0.01157	0.55668
C_Count	0.04852	-0.001183	0.00005	0.55521
Subtype	0.04425	0.0008853	0.00101	NA
O_Count	0.04188	0.006697	0.01231	0.15266
Eads(ASP)	0.03727	-0.004945	-0.01022	0.32674
Eads(1AX8)	0.03631	-0.001947	-0.00555	-0.05016
Mean curvature	0.03037	0.003655	0.00806	-0.32510
Eads(CHL)	0.03007	-0.001938	-0.00308	0.20346
Dipole moment Y	0.0238	0.003027	0.00560	-0.03599
Dipole moment	0.02011	0.0008146	0.00158	0.09336
Dipole moment X	0.01747	-0.0001216	-0.00083	-0.11422

Conclusion

- From Dataset I
  - Experimental setup (dosage, sacrifice day) are key
    - To compare different NMs the exp. procedures must be identical
  - Length of carbonaceous materials (as is the "Atom Count C") is highly positively correlated to the Neutro Ball cell count
    - Long CNTs (NRCWE 45-50) are like 'asbestos' particles
- From Dataset II
  - BET highly correlated (negatively)
  - Length and Diameter (Curvature)
  - Eads DPPC

IF WE™

DASSAULT SYSTEMES