



Article

# Carbon Nanotubes' Effect on Mitochondrial Oxygen Flux Dynamics: Polarography Experimental Study and Machine Learning Models using Star Graph Trace Invariants of Raman Spectra

Michael González-Durruthy<sup>1</sup> , Jose M. Monserrat<sup>1,\*</sup> , Bakhtiyor Rasulev<sup>2</sup> , Gerardo M. Casañola-Martín<sup>3</sup>, José María Barreiro Sorrivas<sup>4</sup>, Sergio Paraíso-Medina<sup>5</sup> , Víctor Maojo<sup>5</sup>, Humberto González-Díaz<sup>6,7</sup>, Alejandro Pazos<sup>8,9</sup> and Cristian R. Munteanu<sup>8,9,\*</sup>

- <sup>1</sup> Institute of Biological Science (ICB), Federal University of Rio Grande, Rio Grande, RS 96270-900, Brazil; gonzalezdurruthy.furg@gmail.com
- <sup>2</sup> Department of Coatings and Polymeric Materials, North Dakota State University (NDSU), Fargo, ND 58102, USA; bakhtiyor.rasulev@ndsu.edu
- <sup>3</sup> Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada; gmaikelc@gmail.com
- <sup>4</sup> Computer Science School (ETSIINF), Polytechnic University of Madrid (UPM), Calle de losCiruelos, Boadilla del Monte, 28660 Madrid, Spain; jmbarreiro@fi.upm.es
- <sup>5</sup> Biomedical Informatics Group, Artificial Intelligence Department, Polytechnic University of Madrid, Calle de los Ciruelos, Boadilla del Monte, 28660 Madrid, Spain; sparaiso@infomed.dia.fi.upm.es (S.P.-M.), vmaajo@fi.upm.es (V.M.)
- <sup>6</sup> Department of Organic Chemistry II, University of the Basque Country UPV/EHU, 48940 Leioa, Biscay, Spain; gonzalezdiazh@yahoo.es
- <sup>7</sup> IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Biscay, Spain
- <sup>8</sup> INIBIC Institute of Biomedical Research, CHUAC, UDC, 15006 Coruña, Spain; apazos@udc.es
- <sup>9</sup> RNASA-IMEDIR, Computer Sciences Faculty, University of Coruña, 15071 Coruña, Spain
- \* Correspondence: josemmonserrat@gmail.com (J.M.M.); c.munteanu@udc.es (C.R.M.); Tel.: +34-981-167-000 (ext. 1302) (C.R.M.)

Received: 7 October 2017; Accepted: 8 November 2017; Published: 11 November 2017

**Abstract:** This study presents the impact of carbon nanotubes (CNTs) on mitochondrial oxygen mass flux ( $J_m$ ) under three experimental conditions. New experimental results and a new methodology are reported for the first time and they are based on CNT Raman spectra star graph transform (spectral moments) and perturbation theory. The experimental measures of  $J_m$  showed that no tested CNT family can inhibit the oxygen consumption profiles of mitochondria. The best model for the prediction of  $J_m$  for other CNTs was provided by random forest using eight features, obtaining test R-squared ( $R^2$ ) of 0.863 and test root-mean-square error (RMSE) of 0.0461. The results demonstrate the capability of encoding CNT information into spectral moments of the Raman star graphs (SG) transform with a potential applicability as predictive tools in nanotechnology and material risk assessments.

**Keywords:** carbon nanotubes; cytotoxicity; mitochondria oxygen mass flux; Raman spectroscopy; graph theory; spectral moments

## 1. Introduction

Carbon nanotubes (CNTs) have attracted great interest for their promising applications in the fields of biomaterials and nano-biotechnology. Therefore, the evaluation of their toxicity in biological systems is a goal of major importance for the biomaterial sciences. Currently, evidence has been

accumulating regarding the CNT toxicity associated with mitochondrial dysfunction and apoptosis [1]. Some in vitro studies have demonstrated that CNTs exhibit cytotoxicity after their accumulation in the mitochondria matrix and/or by affecting the function of mitochondrial respiratory complexes of the inner membrane [2,3].

However, it is still not known which bio-energetic mechanisms (the inhibition of adenosine diphosphate/adenosine triphosphate (ADP/ATP)-transport, uncoupling effects on oxidative phosphorylation, the induction of mitochondrial permeability transition pores, etc.) are responsible for the initiation of CNT mitochondrial damage [4–9]. Particularly, the study of mitochondrial dysfunction based on perturbations of the mitochondrial oxygen mass flux induced by CNTs could be decisive for the in vitro prediction of the no-observed-adverse-effect level (NOAEL) [10] and the evaluation of the selective nanotoxicity (mitochondrial channel nanotoxicity) towards potential biomedical applications in precision medicine. Several scientific reports highlight a major impact of adverse/toxic effects induced by CNT on critical mitochondrial components due to a peculiar mitotropic behavior. Mitochondrial oxygen mass flux regulates the mitochondrial volume linked to the  $\text{Ca}^{2+}$  induction of mitochondrial permeability transition pores (MPTP) and also to the increase of the mitochondrial reactive oxygen species (ROS-levels) based on the partial reduction of molecular oxygen in the mitochondrial respiratory complexes (complex IV). The latter have been extensively characterized and associated with several chronic pathological processes, such as neurodegenerative diseases (Alzheimer, Parkinson, Epilepsy), cardiac ischemia, and cancer. These diseases have currently high levels of morbidity and mortality, and mitochondrial dysfunction based on oxygen mass flux mechanisms has been indirectly or directly involved.

On the other hand, Raman spectroscopy provides information on the chemical fingerprints of molecules, biomolecular systems and nanoscale structures: DNA [11], proteins [12], antibodies [13] and CNTs [14]. In addition, the presence of the G band ( $1580\text{ cm}^{-1}$ ) in the Raman spectra of SWCNTs has been corroborated in mitochondria associated with incipient colloid–osmotic swelling or the induction of mitochondrial permeability transition pores [15]. Recent studies using Raman spectroscopy and polarographic methods have shown that CNTs alter cytochrome c electron transfer and modulate mitochondrial function at a critical concentration of  $10\text{ }\mu\text{g/mL}$  [16]. Previous quantitative structure–activity relationship (QSAR) models of CNTs in terms of mitochondrial respiratory function have been reported [10]. However, there are only few studies about the relationship between a CNT's physicochemical parameters from an oxygen mass flux perspective.

In principle, the Raman spectra signals of different CNTs can be used as inputs for machine learning (ML) methods to predict a dose–effect relationship for the biological properties of CNTs. Nevertheless, the Raman spectra of CNTs have many peaks ( $>1000$  points), making statistical analysis still possible, but somewhat difficult. A possibility for this kind of signal is to compress them into another series of numerical parameters that quantify useful structural information on all the spectra. In a previous work, the star graphs (SG) of the Raman spectra of CNTs [17] were introduced. The idea is to transform the signal into a network with star graph (SG) topology. Next, different invariants can be calculated from the adjacency matrix associated with this SG graph representation of the spectra. Afterwards, the new invariants are supposed to contain useful information, compressed and used as input in ML experiments. Last, as a result of the ML study, predictive models are obtained, which are able to connect the Raman spectra with the biological activity under study. In fact, our group have used this scheme based on SRN transforms to model biological properties from protein sequences [18–20], nucleic acid sequences [21], blood protein mass spectra [22] and electroencephalogram (EEG) signals [23]. These studies have used different matrix invariants such as Markov–Shannon entropies or matrix trace invariants  $Tr_k$  (also known as spectral moments) to compress the information from proteins, gene, etc. In a previous work, the research focused on how to use Markov–Shannon entropies to compress information from Raman spectra [17]. However, there is no report on the use of  $Tr_k$  values in this sense.

This work is aimed at combining experimental and computational techniques to provide a heuristic solution to the above-discussed problem. Firstly, high-resolution respirometry (HRR) polarographic (Oxygraph-2k) assays are used for the first time to address this issue. This methodology provides a quick and reproducible means to measure the rate of oxygen consumption by mitochondria isolated from different tissues using a sensitive Clark-type electrode. Simultaneously, the oxygen mass flux ( $O_2$  flux) in mitochondrial suspensions can be monitored depending on the time of exposure [24–26]. Currently, there are no precedents for this methodology applied to the evaluation of the no-observed-adverse-effect level of CNTs. In addition, a new computational model is developed for the prediction of dose–effect relationships for this property for other CNTs. In doing so, the  $Tr_k$  values of the Raman spectra of CNTs are used as an input to seek a predictive model based on machine learning (ML) and perturbation theory (PT) (PTML model). PTML models are useful to predict the properties of complex molecular systems with simultaneous variations of multiple experimental boundary conditions, such as chemical reactivity, drug metabolism, vaccine peptide epitopes, metabolic networks, and micelle nanoparticles [27–30]. Figure 1 depicts a workflow scheme with the general steps of this work. The current work paves the way for the use of PTML models, polarography, and Raman spectroscopy, and for the experimental and theoretical study of other biological properties of CNTs in the future.

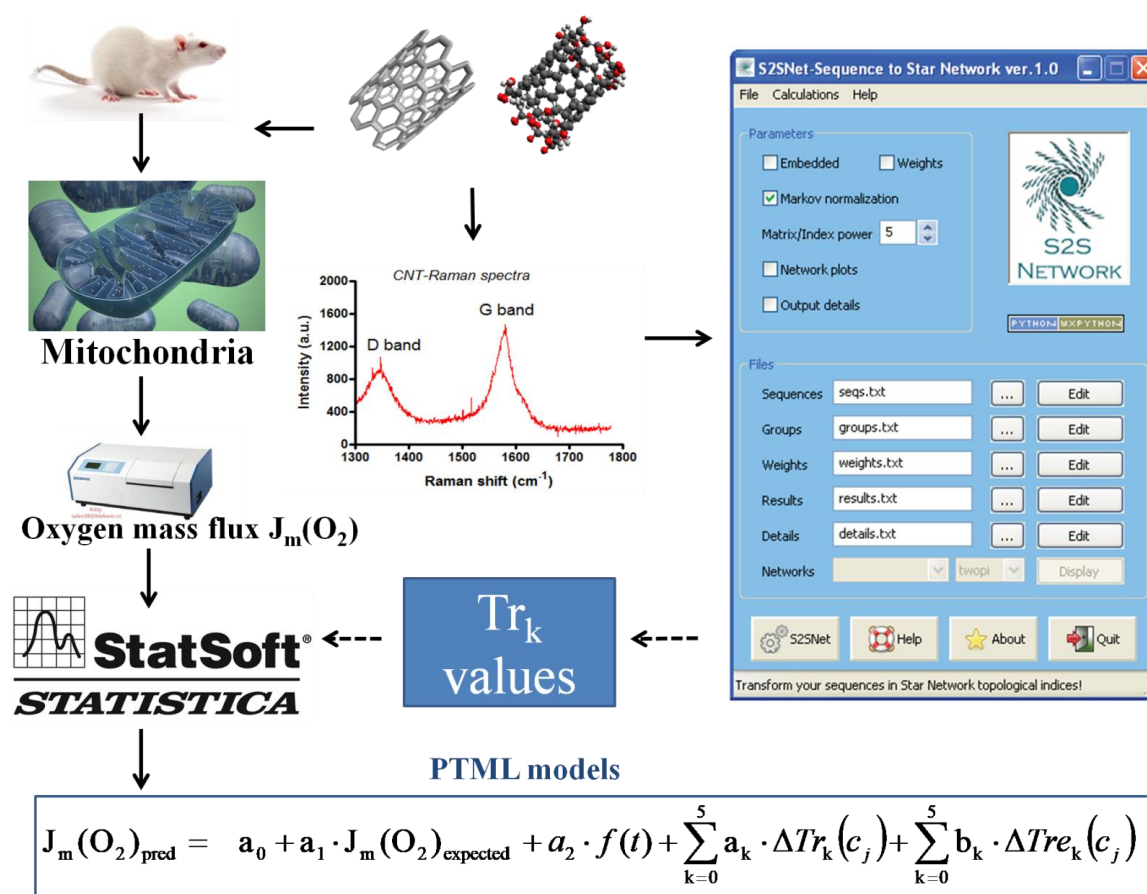


Figure 1. General workflow.

## 2. Results

### 2.1. Experimental Results

High-resolution respirometry (HRR) using Oroboros Instruments (Oxygraph-2k) was applied to evaluate the effects of the CNT family on the bioenergetic mitochondrial function through the

measurement of oxygen mass flux after exposure with different CNTs (from CNT-1 to CNT-9). A default experimental concentration of 5  $\mu\text{g}/\text{mL}$  for all CNTs was established to assess the different contribution of the remaining physical–chemical parameters in the NOAEL for the oxygen mass flux response. Figure 2 shows a representative profile of the mitochondrial oxygen mass flux of isolated rat liver mitochondria, showing the small pulses of ADP-titration associated with transient increments of  $\text{O}_2$  flux during the period of ATP synthesis or V3 state (ADP-dependent respiration) = ratio of mitochondrial ADP-flux ( $J_m(\text{ADP})$ )/( $J_m(\text{O}_2)$ ) + mitochondrial inorganic pyrophosphate Pi-flux ( $J_m(\text{Pi})$ ) for the experimental condition of untreated-rat liver mitochondria (untreated-RLM) and pre-incubation with the different CNTs (RLM + CNTs 5  $\mu\text{g}/\text{mL}$ ). The respiratory substrates (ADP) and the uncoupling agent carbonyl-cyanide p-trifluoromethoxyphenylhydrazone (FCCCP) were added where indicated by the arrow. In addition, in the supplementary information file, a figure that depicts a control profile of the mitochondrial oxygen mass flux of isolated RLM was included.

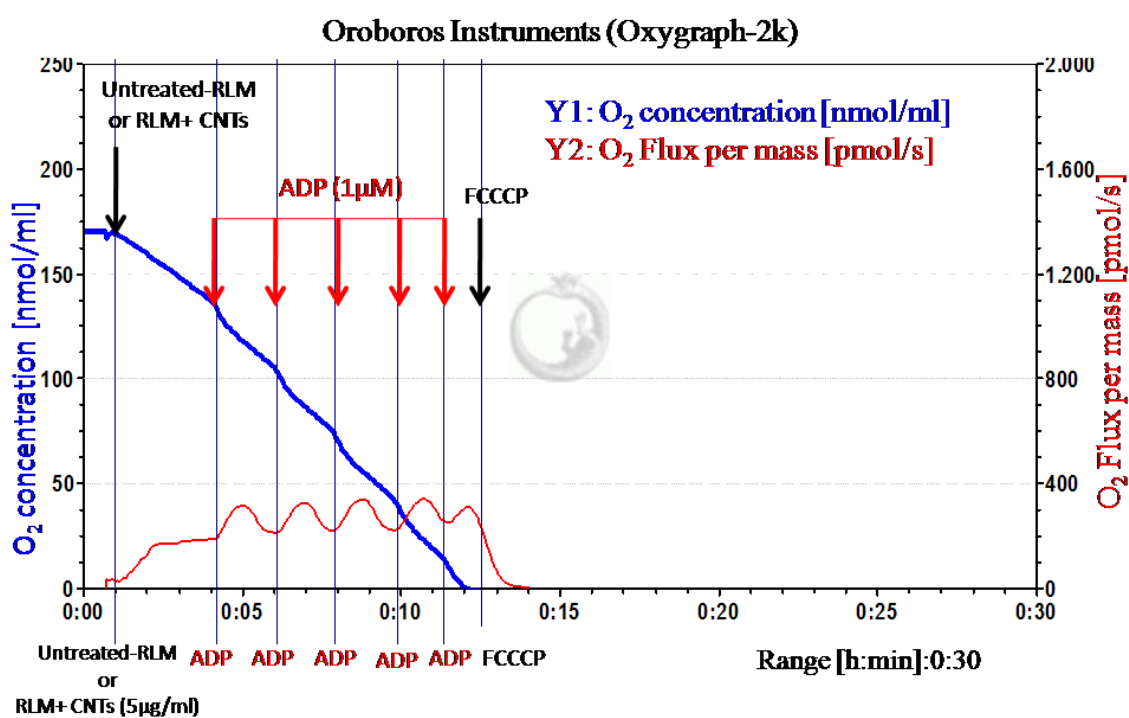
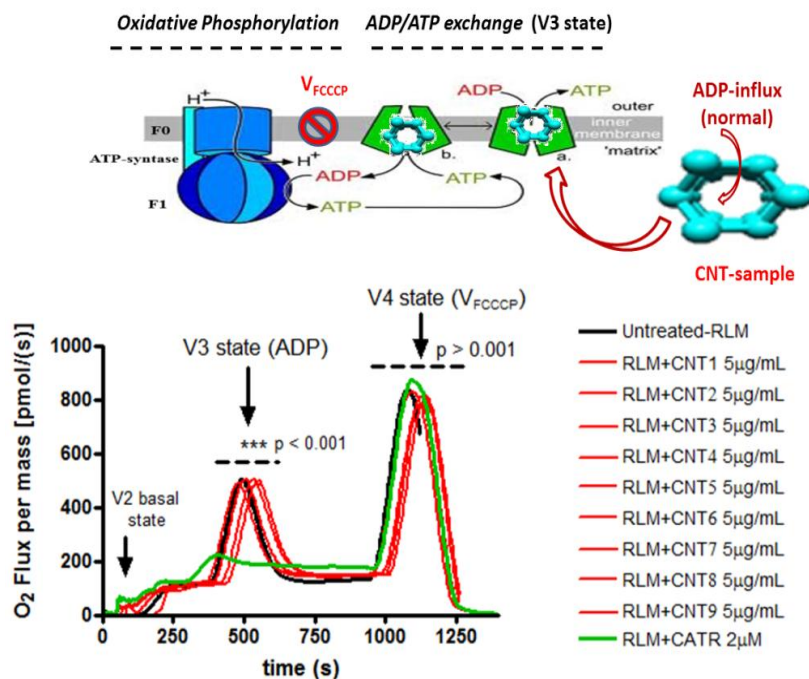


Figure 2. Representative profiles of the mitochondrial oxygen mass flux of isolated rat liver mitochondria (Y2: Red curve).

Figure 3 shows a schematic representation of mitochondrial ADP/ATP exchange and oxidative phosphorylation based on a proposed mechanism linked to the profiles of the mitochondrial oxygen mass flux of isolated rat liver mitochondria. It shows the different states of the mitochondrial respiration V2 state (basal respiration), V3 state (ADP-dependent respiration), V4 (or VFCCCP state of ADP-independent mitochondrial respiration) in untreated rat liver mitochondria or untreated RLM (black curve), CNT-treated RLM at concentration 5  $\mu\text{g}/\text{mL}$  (red curves) and treated RLM with CATR (or ADP-inhibitor) (green curve). The respiratory substrates (ADP) and the uncoupling agent (FCCCP) were added where indicated by the arrow. These results are representative for the three experiments using Oroboros Instruments (Oxygraph-2k). \*  $p < 0.001$  to statistical differences between CATR (or ADP-inhibitor) (green curve) and CNT-treated RLM at concentration 5  $\mu\text{g}/\text{mL}$  (red curves).



**Figure 3.** Representation of mitochondrial ADP/ATP exchange and oxidative phosphorylation. \*\*\*p is used to represent the significant statistical differences between V3 state-ADP-dependent mitochondrial O<sub>2</sub> flux from the RLM + CNT treated groups (CNT1-9) and V3 state-ADP-dependent mitochondrial O<sub>2</sub> flux from the RLM + Carboxyatractyloside (CATR, a specific inhibitor of ADP-mitochondrial transport).

## 2.2. Computational Results

The variation of the mitochondrial oxygen mass flux in the presence of CNTs was modeled using the PTML methodology based on the SG trace invariants of the Raman spectra ( $Tr_{0-5}$ ) of CNT. The RRegrs tool was adapted for calculations on an High Performance Computing (HPC) cluster and was used to test three types of regression methods: Linear Multi-regression (LM), Neural Network (NN) and Random Forest (RF). The objective of the modeling was to find the best prediction model for the mitochondrial oxygen flow in the presence of CNTs. The linear and non-linear methods were used with 10 different random data splits (75% for training–25% for test) using a custom script presented in the Figshare repository (<https://doi.org/10.6084/m9.figshare.3472349>). This script normalizes and splits datasets in a similar way to the RRegrs tool with the parameters  $CVtypes = repeatedcv$  (10-cross-validations) and default number of data splits ( $iSplitTimes = 10$ ). After the normalization and removal of the correlated features (default RRegrs methodology  $cutoff = 0.9$ ), eight input features remain in the final dataset. Thus, the Figshare repository presents several files: “ds2.full.Tr.csv”, the initial dataset without any filters (all features), and “ds2.corr.Tr.csv”, the final dataset after removing correlated features and using normalization. The last dataset was used with the script “CreateNormalisedSplits.R” to create 10 random splits of data as training and test subsets to test them with regression methods.

The first two are the expected values of flux,  $J_m(O_2)_{expected}$  and the duration of the experiment ( $t$ ). The other six input variables are moving average (MA) operators, including at least one MA for each experimental condition. These MA operators are:  $\Delta Tr_0(rep)$ ,  $\Delta Tre_3(rep)$ ,  $\Delta Tre_0(CNTtype)$ ,  $\Delta Tre_4(Func.Type)$ ,  $\Delta Tre_0(Solvent)$ , and  $\Delta Tre_5(Solvent)$ . Please note that the MA operators  $\Delta Tr_k(c_j) = Tr_k - \langle Tr_k(c_j) \rangle$  and  $\Delta Tre_k(c_j) = Tre_k - \langle Tre_k(c_j) \rangle$  quantify the deviation of the  $Tr_k$  or  $Tre_k$  of the Raman spectra of a specific CNT from the expected values  $\langle Tr_k(c_j) \rangle$  or  $\langle Tre_k(c_j) \rangle$ , measured for all Raman spectra and recorded for all CNTs with the specific experimental condition ( $c_j$ ). The symbol  $Tr_k$  refers to traces calculated from linear graphs of the Raman spectra and the symbol  $Tre$  for graphs with recurrence information embedded. Table 1 shows the minimum, maximum, and mean values for R-squared/root mean squared error of training and test subsets ( $R^2_{tr}/RMSE_{tr}$  and  $R^2_{ts}/RMSE_{ts}$ ) for



the eight-feature dataset (10 splits). The first observation in terms of results is that the models were not over-fitted because the differences between the training and test statistics were small. Additional statistics are available online at Figshare [31] (ds2.Tr.models.xlsx).

**Table 1.** Predictive model based on Machine Learning and Perturbation Theory (PTML) statistics for the evaluation of mitochondrial oxygen flow modifications due to CNTs (10 random splits for each method).

Regression Method	Statistics	Training		Test	
		$R^2_{tr}$	RMSE <sub>tr</sub>	$R^2_{ts}$	RMSE <sub>ts</sub>
Linear Multi-regression (LM)	Mean	0.358	0.0959	0.356	0.0954
	Min	0.349	0.0954	0.340	0.0932
	Max	0.363	0.0966	0.384	0.0969
Neural Network (NN)	Mean	0.645	0.0709	0.672	0.0681
	Min	0.626	0.0697	0.620	0.0613
	Max	0.659	0.0727	0.739	0.0738
Random Forest (RF)	Mean	0.855	0.0455	0.856	0.0452
	Min	0.851	0.0451	0.853	0.0431
	Max	0.858	0.0462	0.863	0.0461

The file ds2.Tr.models.xlsx from Figshare presents the statistics for each regression model and individual data split. In addition, minimum, maximum, and average values of the statistics are presented. The final results include the best model from an individual split and its model is saved as an R object. Thus, it is possible to directly use the best model by loading it with an R script and making any prediction. Average values of  $R^2_{ts}$  and RMSE<sub>ts</sub> for LM, NN and RF are presented in the same file.

### 3. Discussion

#### 3.1. Discussion of Experimental Results

The results showed that all the tested CNT family did not inhibit (or affect) the profiles of oxygen mass flux in isolated rat liver mitochondria after the sequential addition of ADP intermittent pulses, which characterize the state V3 of respiration (ADP-dependent) for the untreated RLM and treated RLM with CNTs (5 µg/mL) (see Figure 2, red profile of oxygen flux). Note that for this instance, no significant differences were detected in the profiles of mitochondrial oxygen mass flux, compared with the strong inhibition of the state V3 of respiration detected for mitochondrial treated with carboxyatractyloside a specific inhibitor of ADP-mitochondrial transport (CATR-treated RLM), as shown in Figure 3 (red profile of oxygen flux). Moreover, the oxidative phosphorylation was not affected by ATP synthase, which depends on the ADP transport by ADP/ATP mitochondrial carrier between the cytosol and mitochondrial matrix under physiological normoxic conditions. The in vitro results suggest a non-inhibitory biochemical response of oxygen mass flux in isolated rat liver mitochondria. Furthermore, treated RLM with CNTs maintained the normally-induced uncoupling response of state V4 (or Vf<sub>ccp</sub>) after the addition of FCCCP 2 µM (classical uncoupling agent of the mitochondrial oxidative phosphorylation) according to an increase in the oxygen mass flux state V4 of respiration between 750 and 1250 seconds for all CNTs tested, as shown in Figure 3.

According to these results, several aspects should be considered in order to explain the relevance of CNT NOAEL in terms of mitochondrial oxygen mass flux response. Covalently functionalized CNTs (oxidized-CNTs) and/or with point defects characterized by the D band of Raman spectra (with a characteristic peak at 1350 cm<sup>-1</sup>) are expected to have greater biocompatibility than pristine CNTs [15]. This may be due to the ability of the OH and COOH groups of oxidized CNT or π-bond of the sp<sup>2</sup> of pristine CNT wall to form several adducts with the basal oxygen-free radical [32], released by the mitochondrial complex I and III at between the V2 and V3 state of mitochondrial respiration. In this context, the non-significant respiratory effects from a low CNT concentration (5 µg/mL) could be recognized as a typical pharmacodynamic criterion of NOAEL for CNTs, similar to the sub-clinical effects of traditional lipophilic agents with mitochondrial mechanisms reported in the literature.

As mentioned in the introduction of this work, CNTs can modulate mitochondrial function at a critical concentration of 10  $\mu\text{g}/\text{mL}$  [25].

In this sense, it should be pointed out that the biocompatibility/toxicity relationship of the tested CNT family may be important in predicting the no-observed-adverse-effect level. This is not only limited to considerations of dosimetry in terms of concentration, but also applies to the influence of the physicochemical nanodescriptors of the tested CNTs (Raman nanodescriptors). Thus, the prediction of uncertainty factors or encrypted information in the CNT structure, such as new Raman spectra nanodescriptors, could be used to model the dosimetric criteria anticipated to be without an increased risk for CNT adverse effects [33,34].

### 3.2. Discussion of Computational Results

Discussion of the PTML computational study. LM produced poor results (mean  $R^2_{\text{ts}} = 0.356$ ; mean  $\text{RMSE}_{\text{ts}} = 0.0954$ ), showing that the relationship between the input moving averages (MAs) of the CNT Raman SG spectral moment and the mitochondrial oxygen flow is not a linear one, or that these features did not include enough information to model this relationship. Starting with the non-linear methods, such as NN and RF, the performance of the regression model improved. Thus, NN provided a mean  $R^2_{\text{ts}}$  of 0.672 and a mean  $\text{RMSE}_{\text{ts}}$  of 0.0681. The best NN model (split 3) was a single-hidden layer with 10 neurons, a structure of 8-10-1 (eight inputs, 10 neurons in one hidden layer, one output), a weight decay of 0.001, and an  $R^2_{\text{ts}}$  of 0.739 and an  $\text{RMSE}_{\text{ts}}$  of 0.0613 [31].

The RF predictor significantly improved the regression performance for mitochondrial oxygen flow: a mean  $R^2_{\text{ts}}$  of 0.856 and a mean  $\text{RMSE}_{\text{ts}}$  of 0.0452 were obtained. Compared with the NN statistics,  $R^2_{\text{ts}}$  increased by 0.184 and  $\text{RMSE}_{\text{ts}}$  decreased by 36.6%. The best RF model (split 2) has 50 trees, an  $R^2_{\text{ts}}$  of 0.863 and an  $\text{RMSE}_{\text{ts}}$  of 0.0461. The variation in the RF error with the number of trees is shown in Figure 4. The model could be downloaded from the free online repository [31] and it could be used for future predictions or may be included in future R applications. The use of all initial 34 features had no important improvements in the regression performance with mean values for LM, NN and RF of 0.356, 0.601 and 0.857 (see all details online [31]). This means that the other features obtained no useful information to improve the current model. In order to check the model quality, the regression receiver operator characteristic (RROC) curve [35] for the test subset is shown in Figure 5. The RROC demonstrates the performance of the RF model.

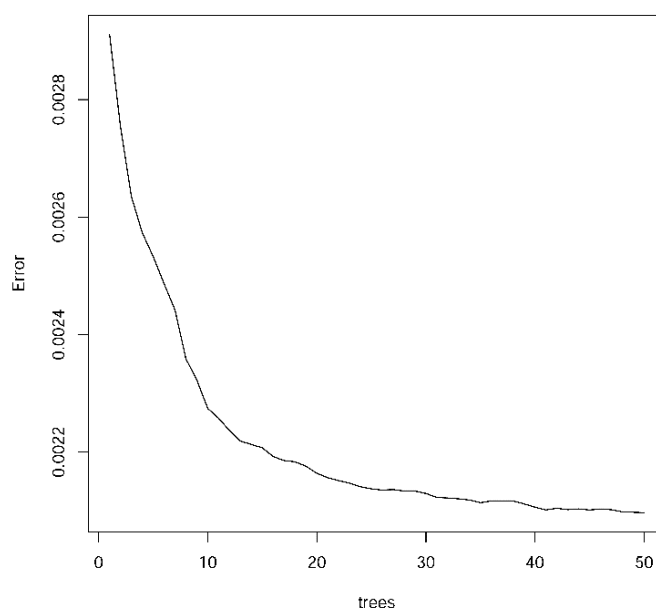
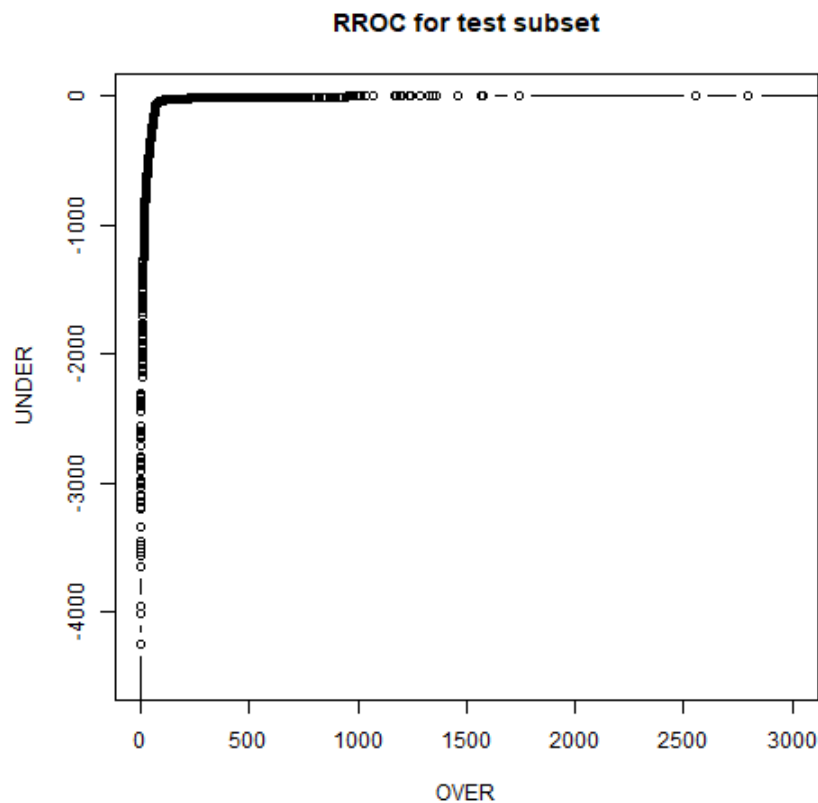


Figure 4. RF error with the number of trees for regression models.



**Figure 5.** Regression receiver operator characteristic (RROC) curves for RF best model (test subset).

#### 4. Materials and Methods

##### 4.1. General Workflow

The aims of this paper are:

- (a) The measurement of the CNT effect on the mitochondrial oxygen mass flux with polarography;
- (b) The definition and calculation of the matrix trace invariants ( $Tr_k$ ) of SG transforms of Raman spectra for a series of CNTs for the first time;
- (c) The use of the  $Tr_k$  values as input to seek new PTML models able to predict CNTs' effect on mitochondrial oxygen mass flux.

Thus, the following steps were taken (see Figure 6):

- (1) Experimental measurements of the mitochondria oxygen mass flux in the presence of different CNT types;
- (2) Transformation of CNT's Raman spectra into SG spectral moments;
- (3) Calculation of the expected values of the mitochondria oxygen mass flux and the moving averages of the SG spectral moments under different experimental conditions;
- (4) Search for the best regression PTML models using the RRegrs package in R (<https://github.com/enanomapper/RRegrs/>).



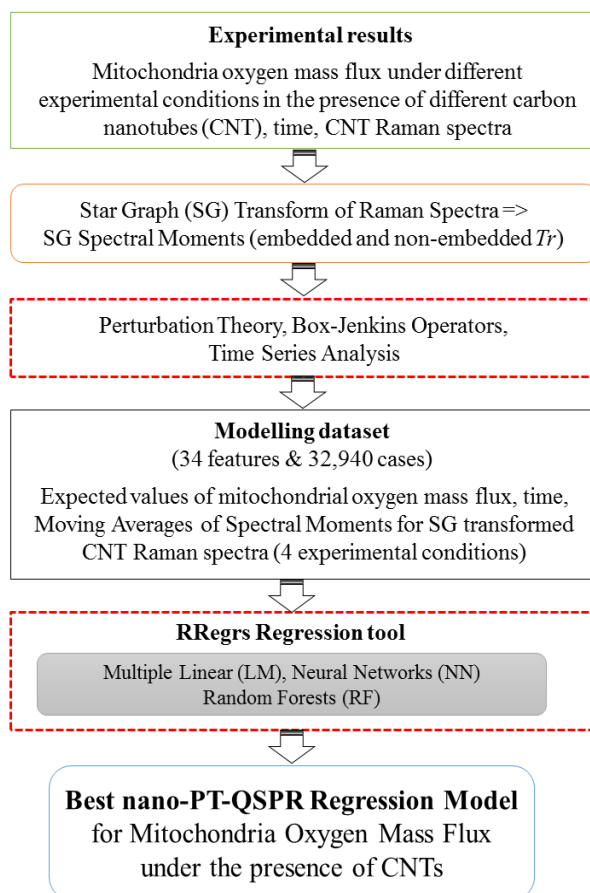


Figure 6. General workflow for the machine learning analysis.

## 4.2. Experimental Methods

### 4.2.1. General Procedures

In the Supplementary Information (SI) file, a detailed description of the following sections is included: sample preparation, reagents and solutions, Raman spectra recording, animal welfare, and isolation of rat liver mitochondria (RLM). The RLM were isolated by standard differential centrifugation according to the experimental procedures established in the literature [36].

### 4.2.2. Monitoring Mitochondrial Oxygen Mass

The HRR method was used, along with Oroboros Instruments, DatLab Version 4.2.1.50 (Oxygraph-2k). This methodology included the use of a 2 mL glass chamber equipped with a magnetic stirrer. Mitochondrial O<sub>2</sub> mass flux (pmol/seg) in the absence and presence of different carbon nanotubes (CNT-1 to CNT-9) was monitored and measured as the negative time derivative of an oxygen concentration (nmol/mL). The mitochondrial oxygen mass flux values were corrected for the small amount of back diffusion of oxygen from materials within the chamber, any leak of oxygen from outside the vessel, and oxygen consumed by the polarographic electrode [26,37,38].

With this in mind, the RLM isolated (1 mg protein/mL) were energized with 5 mM potassium succinate (plus 2.5 μM rotenone) in a standard incubation medium, consisting of 125 mM sucrose, 65 mM KCl, 2 mM inorganic phosphate (K<sub>2</sub>HPO<sub>4</sub>) and 10 mM potassium hydroxide-2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES-KOH) pH 7.4 at 20 °C in a standard respiration medium. The experimental approach was calibrated using the oxygen content of an air saturated medium [38]. All the aforementioned steps were performed by the pre-incubation of

isolated rat liver mitochondria with 5  $\mu\text{g}/\text{mL}$  for all the CNT-treated groups. This level of concentration was considered based on the NOAEL criteria mentioned in the introduction of this work. For this instance, 5  $\mu\text{g}/\text{mL}$  of the CNT concentration is the corrected half value of the CNT concentration (10  $\mu\text{g}/\text{mL}$ ) used by Ma et al., who promoted an incipient colloid–osmotic swelling or low induction of mitochondrial permeability transition pores, noticeably detected by Raman spectroscopy [15,16]. The total number of the collected data points was 32,940. The data obtained from the mitochondrial oxygen mass study in the presence of CNTs and the CNT spectra were used to search for a theoretical model which predicts mitochondrial oxygen mass in the presence of new CNTs.

### 4.3. Computational Methods

#### 4.3.1. Trace Invariants of Raman Spectra

One of the objectives of this work is to develop a mathematical model able to predict the biological effect of CNTs using as an input the information extracted from Raman spectra. Thus, a new type of parameters is proposed, calculated by the application of a Star Graph (SG) transform to the Raman spectra. The SRN transform method, which has been recently introduced and published by our group [17], uses graphs and network theory tools, and is different from a classic Fourier transformation. The transform technique used herein converts the Raman spectra values into sequences of characters and creates the corresponding SG of this signal. The SG of any sequence/signal may be constructed using the S2SNet tool [39]. To construct the SG transform of a Raman spectrum, the latter should be split into intervals of 100 units, from 0 to 1800. As a result, the maximum number of SG branches is 18 and corresponds to characters from “a” to “r”. Subsequently, the adjacency matrix  $\mathbf{A}$  should be constructed for this sequence of characters (spectral sequence). The matrix  $\mathbf{A}$  was enriched by adding the information about the recurrence to the same type of term in the spectral sequence. After this step, the matrix  $\mathbf{A}_e$  was obtained, with recurrence information embedded (e). As a result, different invariants from the matrices  $\mathbf{A}$  and  $\mathbf{A}_e$  can be calculated. In this work, the matrix trace values ( $Tr_k$  and  $Tre_k$ ;  $k = 0-5$ ) were calculated, also known as the spectral moments [39] of the matrices  $\mathbf{A}$  and  $\mathbf{A}_e$  obtained after the SG transformation of the Raman spectra. Our hypothesis is that these spectral moments encode useful structural information that can be responsible for biological activity and further predictive studies (see Figure 1).

#### 4.3.2. PTML Model

The current section describes the application of the algorithm of PTML heuristic models [30] in order to study the effect of different CNTs on the mitochondrial oxygen mass flux under different experimental conditions. The general equation of a linear PTML heuristic model could be described by Equation (1).

$$J_m(O_2)_{pred} = a_0 + a_1 \cdot J_m(O_2)_{expected} + a_2 \cdot f(t) + \sum_{k=0}^5 b_k \cdot \Delta Tr_k(c_j) + \sum_{k=0}^5 c_k \cdot \Delta Tre_k(c_j) \quad (1)$$

where  $J_m(O_2)_{pred}$  is the predicted mitochondrial oxygen mass flux. The term  $J_m(O_2)_{expected} = \langle J_m(O_2) \rangle$  is the average value of  $J_m(O_2)$  for different subsets of experimental conditions (expected value of  $J_m(O_2)$ ). The other input values  $\Delta Tr_k$  and  $\Delta Tre_k$  are the moving average (MA) operators. The MA operators  $\Delta Tr_k(c_j) = Tr_k - \langle Tr_k(c_j) \rangle$  and  $\Delta Tre_k(c_j) = Tre_k - \langle Tre_k(c_j) \rangle$  quantify the deviation of  $Tr_k$  or  $Tre_k$  of the Raman spectra for a specific CNT from the expected average values  $\langle Tr_k(c_j) \rangle$  or  $\langle Tre_k(c_j) \rangle$  measured for all Raman spectra, recorded for all CNTs with the specific experimental condition ( $c_j$ ). The symbols  $Tr_k$  refer to traces calculated from linear graphs of the Raman spectra and the symbol  $Tre_k$  for graphs with recurrence information embedded. The coefficients  $a_k$ ,  $b_k$ , and  $c_k$  are the linear coefficients of the equation.

### 4.3.3. Model Dataset

The experimental data for the mitochondrial oxygen mass flux ( $J_m$ ) in the presence of CNTs are available as a FigShare repository [31]: 32,940 cases of  $J_m(O_2)$ , CNT1-9, CNTtype, Replicate (rep), Function\_type (chemical modification of CNT), Solvent and time ( $t$ ).  $J_m(O_2)$  was measured under four types of experimental conditions (c) such as Replicate (rep), CNTtype, Functiontype (chemical modification of CNT), and Solvent. Replicate (rep) has two values: 0 or 1, for non-replicated and replicated experiment. CNTtypes are multi-walled carbon nanotubes (MWCNT), mixed single walled/double-walled carbon nanotubes (SW+DWCNT), and single-walled carbon nanotube (SWCNT). CNT type is "0" when the assay is a control assay with a blank solution with a CNT concentration equal to 0. The values of the solvent condition are H<sub>2</sub>O and dimethyl sulfoxide (DMSO). The CNT Function types could have three values: 0 (none), COOH and OH. The average of the SG spectral moments under the experimental conditions are presented on the FigShare platform [31]. The final dataset, used to find the best prediction model is made up of 32,940 cases and 34 input features.

### 4.3.4. PTML Regression Predictors

The raw dataset was normalized and the training and test sub-sets were obtained using 10 splits: 75% training sets (train) and 25% test sets (test) using an R script available online [40]. The regression PTML models were searched with RRegrs, an R integrated framework that provides ten linear and non-linear regression models [41,42]. The selection of the models used the criteria of the  $R_{ts}$  values (regression coefficient for test subset) and the RMSE<sub>ts</sub> (root-mean-square error) corresponding values.

Three types of regression methods of RRegrs were used: multiple linear regression (LM), neural networks regression (NN) [43], and random forest (RF) [44]. Thus, the correlated features were removed using the parameters of RRegrs. For NN and RF, a study of the method parameters was performed. A modified version of RRegrs (batchRRegrs: <https://github.com/cafernandezlo/batchRRegrs>) was used on the BioCAI HPC platform from the University of A Coruña (A Coruña, Spain). The batchRRegrs default values of parameters were generally employed (<https://github.com/cafernandezlo/batchRRegrs/blob/master/batchRRegrs/batchRRegrs.R>). The default optimizations used the *tuneGrid* parameter of the caret training method:

- *NNreg* function used a grid for 200, 300 and 400 neurons in the hidden layer (*.size*) and a decay of 0, 0.01, 0.2, 0.1 (*.decay*) (*method = 'nnet'*);
- *RFreg* function used 1500 trees (*ntree = 1500* for *method = 'rf'*).

An additional number of neurons were tested for NN (1, 5, 10 and 15) and different decay values (0.001 and 0.005). Moreover, an additional number of trees in RF were tested: 5, 10, 20, 30, 40, 50, 100, and 500. The best RF model should use the lowest number of trees and best statistics. The results presented for NN and RF models the parameters for the best models.

The criteria to find the best model apply the RRegrs methodology: maximum  $R_{ts}^2$  and minimum RMSE<sub>ts</sub>. The plots were obtained with custom R scripts. The best regression model which predicts mitochondria oxygen mass flux in the presence of CNTs is available online [31] in order to be used for future predictions. The regression receiver operator characteristic (RROC) curve was constructed using the algorithm from reference [35]. The RROC presented the over-estimation (OVER) against the under-estimation (UNDER). Thus, the curve was drawn by adjusting a shift (a constant that was added or subtracted) for the predictions. This shift is similar to the threshold in the case of classifications.

## 5. Conclusions

The current study presented a mixture of experimental and predictive methodologies to study the effect of different CNTs on the mitochondrial oxygen mass flux. The experimental results showing non-significant respiratory effects from low CNT concentrations (5 µg/mL) could be recognized as a typical pharmacodynamics criterion of NOAEL for CNTs. In this context, the information encrypted in

the Raman spectra of CNT structures can be used as novel nanodescriptors to model the complexity of the dosimetric criteria, as no adverse mitochondrial respiratory effect level (or normal O<sub>2</sub> flux) was found. The R object model and an R script are available online at <https://dx.doi.org/10.6084/m9.figshare.3545561>.

These results show that the SG transform of CNT Raman spectra contains important information, as new CNT nanodescriptors can be combined to provide a prediction model under the experimental conditions over time for the mitochondria oxygen mass flux under the presence of specific CNTs. These in silico results indicate that this methodology can be employed for massive, virtual-based, raw data for Raman spectroscopy in order to make regulatory decisions in the biomaterial sciences.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2079-4991/7/11/386/s1>, Table S1. Properties of CNT families, Figure S1. Raman spectra for carbon nanotubes used in the present study, Figure S2. Control profile of mitochondrial oxygen mass flux of isolated rat liver mitochondria (Y2: Red curve).

**Acknowledgments:** Michael González-Durruthy acknowledges a doctoral fellowship (Postgraduate Students' Program PEC-PG No. 062/2013) from the Brazilian CAPES-CNPq Agencies. This work is supported by the "Collaborative Project on Medical Informatics (CIMED)" funded by the Carlos III Health Institute from the Spanish National plan for Scientific and Technical Research and Innovation 2013–2016 and the European Regional Development Funds (FEDER) [grant numbers PI13/02020, PI13/00280]. In addition, the work was supported by the Galician Network of Drugs R+D REGID (Xunta de Galicia R2014/025), the General Directorate of Culture, Education and University Management of Xunta de Galicia (Ref. GRC2014/049) and the Galician Network for Colorectal Cancer Research (REGICC) (Ref. R2014/039). Finally, the project was supported by the Spanish Ministry of Economy and Competitiveness (MINECO) research grants (UNLC08-1E-002, UNLC13-13-3503 and CTQ2016-74881-P) and the Basque Government (IT1045-16), Ikerbasque, Basque Foundation for Science. Jose M. Monserrat acknowledges the support of Brazilian Agency CNPq (process number 454332/2014-9) and a scientific productivity fellowship (process number 308539/2016-8).

**Author Contributions:** Michael González-Durruthy, Jose M. Monserrat, Bakhtiyor Rasulev, and Gerardo M. Casañola-Martín performed the experiments and analyzed the experimental results. José María Barreiro Sorriveras, Sergio Paraíso-Medina, Víctor Maojo, Humberto González-Díaz, Alejandro Pazos and Cristian R. Munteanu performed the computational study and analyzed the machine learning results. All authors wrote the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Prato, M.; Kostarelos, K.; Bianco, A. Functionalized carbon nanotubes in drug design and discovery. *Acc. Chem. Res.* **2008**, *41*, 60–68. [[CrossRef](#)] [[PubMed](#)]
2. Toropov, A.A.; Leszczynska, D.; Leszczynski, J. Predicting water solubility and octanol water partition coefficient for carbon nanotubes based on the chiral vector. *Comput. Biol. Chem.* **2007**, *31*, 127–128. [[CrossRef](#)] [[PubMed](#)]
3. Ye, S.F.; Wu, Y.H.; Hou, Z.Q.; Zhang, Q.Q. Ros and nf-kappab are involved in upregulation of il-8 in a549 cells exposed to multi-walled carbon nanotubes. *Biochem. Biophys. Res. Commun.* **2009**, *379*, 643–648. [[CrossRef](#)] [[PubMed](#)]
4. Shvedova, A.A.; Castranova, V.; Kisin, E.R.; Schwegler-Berry, D.; Murray, A.R.; Gandelsman, V.Z.; Maynard, A.; Baron, P. Exposure to carbon nanotube material: Assessment of nanotube cytotoxicity using human keratinocyte cells. *J. Toxicol. Environ. Health A* **2003**, *66*, 1909–1926. [[CrossRef](#)] [[PubMed](#)]
5. Bottini, M.; Bruckner, S.; Nika, K.; Bottini, N.; Bellucci, S.; Magrini, A.; Bergamaschi, A.; Mustelin, T. Multi-walled carbon nanotubes induce t lymphocyte apoptosis. *Toxicol. Lett.* **2006**, *160*, 121–126. [[CrossRef](#)] [[PubMed](#)]
6. Cui, D.; Tian, F.; Ozkan, C.S.; Wang, M.; Gao, H. Effect of single wall carbon nanotubes on human HEK293 cells. *Toxicol. Lett.* **2005**, *155*, 73–85. [[CrossRef](#)] [[PubMed](#)]
7. Muller, J.; Huaux, F.; Moreau, N.; Misson, P.; Heilier, J.F.; Delos, M.; Arras, M.; Fonseca, A.; Nagy, J.B.; Lison, D. Respiratory toxicity of multi-wall carbon nanotubes. *Toxicol. Appl. Pharmacol.* **2005**, *207*, 221–231. [[CrossRef](#)] [[PubMed](#)]
8. Heller, D.A.; Baik, S.; Eurell, T.E.; Strano, M.S. Single-walled carbon nanotube spectroscopy in live cells: Towards long-term labels and optical sensors. *Adv. Mater.* **2005**, *17*, 2793–2799. [[CrossRef](#)]
9. Smart, S.K.; Cassady, A.I.; Lu, G.Q.; Martin, D.J. The biocompatibility of carbon nanotubes. *Carbon* **2006**, *44*, 1034–1047. [[CrossRef](#)]

10. Toropov, A.A.; Toropova, A.P.; Pizzo, F.; Lombardo, A.; Gadaleta, D.; Benfenati, E. Coral: Model for no observed adverse effect level (noael). *Mol. Divers.* **2015**, *19*, 563–575. [[CrossRef](#)] [[PubMed](#)]
11. Urabe, H.; Sugawara, Y.; Ataka, M.; Rupprecht, A. Low-frequency raman spectra of lysozyme crystals and oriented DNA films: Dynamics of crystal water. *Biophys. J.* **1998**, *74*, 1533–1540. [[CrossRef](#)]
12. Chou, K.C. Identification of low-frequency modes in protein molecules. *Biochem. J.* **1983**, *215*, 465–469. [[CrossRef](#)] [[PubMed](#)]
13. Piacenza, M.; Grimme, S. Van der waals interactions in aromatic systems: Structure and energetics of dimers and trimers of pyridine. *ChemPhysChem* **2005**, *6*, 1554–1558. [[CrossRef](#)] [[PubMed](#)]
14. Hennrich, F.; Krupke, R.; Lebedkin, S.; Arnold, K.; Fischer, R.; Resasco, D.E.; Kappes, M.M. Raman spectroscopy of individual single-walled carbon nanotubes from various sources. *J. Phys. Chem. B* **2005**, *109*, 10567–10573. [[CrossRef](#)] [[PubMed](#)]
15. Ma, X.; Zhang, L.H.; Wang, L.R.; Xue, X.; Sun, J.H.; Wu, Y.; Zou, G.; Wu, X.; Wang, P.C.; Wamer, W.G.; et al. Single-walled carbon nanotubes alter cytochrome c electron transfer and modulate mitochondrial function. *ACS Nano* **2012**, *6*, 10486–10496. [[CrossRef](#)] [[PubMed](#)]
16. Yang, Z.; Zhang, Y.; Yang, Y.; Sun, L.; Han, D.; Li, H.; Wang, C. Pharmacological and toxicological target organelles and safe use of single-walled carbon nanotubes as drug carriers in treating alzheimer disease. *Nanomed. Nanotechnol. Biol. Med.* **2010**, *6*, 427–441. [[CrossRef](#)] [[PubMed](#)]
17. Gonzalez-Durruthy, M.; Alberici, L.C.; Curti, C.; Naal, Z.; Atique-Sawazaki, D.T.; Vazquez-Naya, J.M.; Gonzalez-Diaz, H.; Munteanu, C.R. Experimental-computational study of carbon nanotube effects on mitochondrial respiration: In silico nano-qspr machine learning models based on new raman spectra transform with markov-shannon entropy invariants. *J. Chem. Inf. Model.* **2017**, *57*, 1029–1044. [[CrossRef](#)] [[PubMed](#)]
18. Fernandez-Lozano, C.; Cuinas, R.F.; Seoane, J.A.; Fernandez-Blanco, E.; Dorado, J.; Munteanu, C.R. Classification of signaling proteins based on molecular star graph descriptors using machine learning models. *J. Theor. Biol.* **2015**, *384*, 50–58. [[CrossRef](#)] [[PubMed](#)]
19. Fernandez-Lozano, C.; Gestal, M.; Pedreira-Souto, N.; Postelnicu, L.; Dorado, J.; Munteanu, C.R. Kernel-based feature selection techniques for transport proteins based on star graph topological indices. *Curr. Top. Med. Chem.* **2013**, *13*, 1681–1691. [[CrossRef](#)] [[PubMed](#)]
20. Fernandez-Blanco, E.; Aguiar-Pulido, V.; Munteanu, C.R.; Dorado, J. Random forest classification based on star graph topological indices for antioxidant proteins. *J. Theor. Biol.* **2012**, *317*, 331–337. [[CrossRef](#)] [[PubMed](#)]
21. Perez-Bello, A.; Munteanu, C.R.; Ubeira, F.M.; De Magalhaes, A.L.; Uriarte, E.; Gonzalez-Diaz, H. Alignment-free prediction of mycobacterial DNA promoters based on pseudo-folding lattice network or star-graph topological indices. *J. Theor. Biol.* **2009**, *256*, 458–466. [[CrossRef](#)] [[PubMed](#)]
22. Vázquez, J.M.; Aguiar, V.; Seoane, J.A.; Freire, A.; Serantes, J.A.; Dorado, J.; Pazos, A.; Munteanu, C.R. Star graphs of protein sequences and proteome mass spectra in cancer prediction. *Curr. Proteom.* **2009**, *6*, 275–288. [[CrossRef](#)]
23. Fernandez-Blanco, E.; Rivero, D.; Rabunal, J.; Dorado, J.; Pazos, A.; Munteanu, C.R. Automatic seizure detection based on star graph topological indices. *J. Neurosci. Methods* **2012**, *209*, 410–419. [[CrossRef](#)] [[PubMed](#)]
24. Wang, X.; Guo, J.; Chen, T.; Nie, H.; Wang, H.; Zang, J.; Cui, X.; Jia, G. Multi-walled carbon nanotubes induce apoptosis via mitochondrial pathway and scavenger receptor. *Toxicol. In Vitro* **2012**, *26*, 799–806. [[CrossRef](#)] [[PubMed](#)]
25. Knobeloch, L.M.; Blondin, G.A.; Harkin, J.M. Use of submitochondrial particles for prediction of chemical toxicity in man. *Bull. Environ. Contam. Toxicol.* **1990**, *44*, 661–668. [[CrossRef](#)] [[PubMed](#)]
26. Horan, M.P.; Pichaud, N.; Ballard, J.W. Review: Quantifying mitochondrial dysfunction in complex diseases of aging. *J. Gerontol. A Biol. Sci. Med. Sci.* **2012**, *67*, 1022–1035. [[CrossRef](#)] [[PubMed](#)]
27. Gonzalez-Diaz, H.; Arrasate, S.; Juan, A.G.; Sotomayor, N.; Lete, E.; Speck-Planche, A.; Ruso, J.M.; Luan, F.; Cordeiro, M.N. Matrix trace operators: From spectral moments of molecular graphs and complex networks to perturbations in synthetic reactions, micelle nanoparticles, and drug adme processes. *Curr. Drug Metab.* **2014**, *15*, 470–488. [[CrossRef](#)] [[PubMed](#)]



28. Vergara-Galicia, J.; Prado-Prado, F.J.; Gonzalez-Diaz, H. Galvez-markov network transferability indices: Review of classic theory and new model for perturbations in metabolic reactions. *Curr. Drug Metab.* **2014**, *15*, 557–564. [[CrossRef](#)] [[PubMed](#)]
29. Gonzalez-Diaz, H.; Perez-Montoto, L.G.; Ubeira, F.M. Model for vaccine design by prediction of B-epitopes of iedb given perturbations in peptide sequence, in vivo process, experimental techniques, and source or host organisms. *J. Immunol. Res.* **2014**, *2014*, 768515. [[CrossRef](#)] [[PubMed](#)]
30. Gonzalez-Diaz, H.; Arrasate, S.; Gomez-SanJuan, A.; Sotomayor, N.; Lete, E.; Besada-Porto, L.; Ruso, J.M. General theory for multiple input-output perturbations in complex molecular systems. 1. Linear qspr electronegativity models in physical, organic, and medicinal chemistry. *Curr. Top. Med. Chem.* **2013**, *13*, 1713–1741. [[CrossRef](#)] [[PubMed](#)]
31. Munteanu, C.R. Regression Models to Predict Mitochondria Oxygen Mass Flux in the preseNce of the Cnts. Figshare, 2016. Available online: <https://dx.Doi.Org/10.6084/m9.Figshare.3545561> (accessed on 7 August 2016).
32. Martínez, A.; Francisco-Marquez, M.; Galano, A. Effect of different functional groups on the free radical scavenging capability of single-walled carbon nanotubes. *J. Phys. Chem. C* **2010**, *114*, 14734–14739. [[CrossRef](#)]
33. Rivera Gil, P.; Oberdorster, G.; Elder, A.; Puentes, V.; Parak, W.J. Correlating physico-chemical with toxicological properties of nanoparticles: The present and the future. *ACS Nano* **2010**, *4*, 5527–5531. [[CrossRef](#)] [[PubMed](#)]
34. Hinderliter, P.M.; Minard, K.R.; Orr, G.; Chrisler, W.B.; Thrall, B.D.; Pounds, J.G.; Teeguarden, J.G. Isdd: A computational model of particle sedimentation, diffusion and target cell dosimetry for in vitro toxicity studies. *Part. Fibre Toxicol.* **2010**, *7*, 36. [[CrossRef](#)] [[PubMed](#)]
35. Hernández-Orallo, J. Roc curves for regression. *Pattern Recogn.* **2013**, *46*, 3395–3411. [[CrossRef](#)]
36. Pedersen, P.L.; Greenawalt, J.W.; Reynafarje, B.; Hullihen, J.; Decker, G.L.; Soper, J.W.; Bustamente, E. Preparation and characterization of mitochondria and submitochondrial particles of rat liver and liver-derived tissues. *Methods Cell Biol.* **1978**, *20*, 411–481. [[PubMed](#)]
37. Lanza, I.R.; Nair, K.S. Functional assessment of isolated mitochondria in vitro. *Methods Enzymol.* **2009**, *457*, 349–372. [[PubMed](#)]
38. Gnaiger, E. Bioenergetics at low oxygen: Dependence of respiration and phosphorylation on oxygen and adenosine diphosphate supply. *Respir. Physiol.* **2001**, *128*, 277–297. [[CrossRef](#)]
39. Munteanu, C.R.; Magalhaes, A.L.; Duardo-Sanchez, A.; Pazos, A.; Gonzalez-Diaz, H. S2Snet: A tool for transforming characters and numeric sequences into star network topological indices in chemoinformatics, bioinformatics, biomedical, and social-legal sciences. *Curr. Bioinf.* **2013**, *8*, 429–437. [[CrossRef](#)]
40. Munteanu, C.R. Cnt Raman for Mitochondrial Oxygen Consumption Data Set and Regression Models. Figshare. Available online: <https://dx.Doi.Org/10.6084/m9.Figshare.3472349> (accessed on 31 January 2017).
41. Tsiliki, G.; Munteanu, C.R.; Seoane, J.A.; Fernandez-Lozano, C.; Sarimveis, H.; Willighagen, E.L. Rregrs: An r package for computer-aided model selection with multiple regression models. *J. Cheminform.* **2015**, *7*, 46. [[CrossRef](#)] [[PubMed](#)]
42. Tsiliki, G.; Munteanu, C.R.; Seoane, J.; Fernandez-Lozano, C.; Sarimveis, H.; Willighagen, E. Using the rregrs R package for automating predictive modelling. In Proceedings of the MOL2NET International Conference on Multidisciplinary Sciences, Bizkaia, Spain, 5–15 December 2015; Volume 1, p. F009.
43. Bishop, C.M. *Neural Networks for Pattern Recognition*; Oxford University Press: New York, NY, USA, 1995.
44. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]

