EMMC-CSA

European Materials Modelling Council

Report

Workshop on physics-based vs data-driven models

June 11-13, 2018

Ångström Laboratory, Uppsala University, Uppsala, Sweden

TABLE OF CONTENT

1. SUMMARY 2
   1.1. Description of the workshop 2
   1.2. Objectives 2
   1.3. Major outcome and Conclusions 2
2. REPORT OF MAIN ACTIVITIES 4
   Session 1 - Data and/or physics I? 4
   Session 2 - Insight from electrons 4
   Session 3 - Coupling & linking ("Multiscale modelling") 5
   Session 4 - Data and/or physics II? 6
   Session 5 - Industrial perspectives 6
   Session 6 - Data and/or physics III? 7
   Session 7 - Towards force-fields for complex systems 7
   Session 8 - ...and the future 8
3. GENERAL DISCUSSION FEEDBACK 9
4. PROGRAMME 10

This event has received funding via the EMMC-CSA project from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 723867

1. Summary

1.1. Description of the workshop

The workshop on "physics-based vs data-driven models" was run inside the conference "Multi-scale modelling (Coupling and linking) of materials and molecules in the 21st century: Physics-based or data-driven?" which was held June11-13, 2018 at the Ångström Laboratory of Uppsala University.

The conference had the purpose to bring together two scientific communities: the more traditional modelers using and developing physics-based methods and models and those applying and creating data-based informatics models now more and more rapidly entering into materials modelling. As they represent two different schools but still solving similar problems it created continuous debate on the pros and cons of both methods but also an agreement that they will both develop together and complement each other, especially, when more and more reliable data appear and become accessible to the community. The current status of data was a frequently touched topic; should all data be open to everybody or not? How should it be organized? Do we need ontologies? Also an open EMMC session was included to discuss this current report for the European Commission.

As far as we are aware, this was the first meeting of its kind anywhere, i.e. a meeting with exactly this focus. The interest from industry was significant with participation from DCS Computing, DSM (a Dutch multinational company in the fields of health, nutrition and materials), Goodyear Dunlop Tires Operations SA (Luxembourg), Virtual Lab (Korea), Granta design (UK), Materials Design S.A.R.L (France), Solvay S.A. (big chemical company, France), Swerea KIMAB AB, Unilever Research, ACCESS e.V., Dow Benelux B.V., Noruna AB. Still on the non-academic side, the following industry-active institutes were represented: Teknova AS (Norway), Korea Institute of Energy Research (KIER), Korea Institute of Science and Technology (KIST), Helmholtz-Zentrum Geestacht and more. There were over a hundred participants from 16 countries. The audience was composed of junior and senior experts in academia and in industry as well as software developers and translators. Specialised presentations as well as expert-talks of a more tutorial character were included. Every scientific session ended with a summary of the highlights raised by speakers and audience collected by the chair-persons.

It turned out to be clever to use the slightly provocative word *versus* in the subtitle, rather than *and.* Most speakers felt a need to comment on it. Overall, the audience and speakers agreed that there is a need for both branches in synergy. At the same time it was clear that this is a fresh topic where there is a lot of interest, and confusion, and a great mission to be achieved.

1.2. Objectives

- To start a focused discussion about the rapidly changing landscape in modelling and simulations in the materials and molecular sciences: expectations and challenges and the consequences it will have on processes, decisions and R&D in industry and academia and for the roles of MAN, MOD, SWO and TRANSLATOR stakeholders.
- To formulate the recommendations for the continued EMMC efforts concerning the identification of model & workflow gaps and lack of information between physics- and data-oriented materials modelling and how to make the best use of them both in complementing each other, based on input from the presenters and the open discussion with all participants.

1.3. Major outcome and Conclusions

A major outcome of the meeting are the following recommendations.

Recommendation 1: There are roughly three categories of materials modelers & method/model developers.
(i) Natural sciences: Actors (in academia) developing/applying the state-of-the-art methods not worrying about the

computing cost or otherwise the time it takes the get results, only that they are accurate.
(ii) *Engineering sciences*: developing phenomenological models and applying thermodynamics and equations-of-state, nearly all parameterized on experiments and
(iii) *Industrial users* who need the results fast and they need to be reasonably accurate. Most often the right trends are enough to give some predictive power. They are generally not interested in the state-of-the-art methods unless they are fast. *How do we reconcile these different needs*?
Academia should be urged to, and obtain the means to, make a focused effort on developing fast -and robust- models and workflows combining traditional and new approaches. We recommend a special European level call as an instrument to develop new hybrid schemes to combine with electronic, atomistic, mesoscopic and continuum models with ML/data-based schemes directed towards key materials-based industries resulting fast and robust tools.

Recommendation 2: Data-based modelling is characteristically highly multidisciplinary but firmly rooted to mathematical statistics and applicable on every discipline producing and analyzing numerical, textual, geometric, visual or audible data all requiring their own flavors. This has resulted in a large number of branching machine learning methods. We suggest an action, similar to EMMC, to
(i) make an inventory of ML methods applied or applicable in materials science and design
(ii) classify the methods (preferably using an ontology) and
(iii) start a program to promote the methods,
(iv) educate the materials modelling community of their benefits and how to optimally combine them with particle/continuum based modelling.

Recommendation 3: Data-based modelling and its value depend intimately on reliable data. The data bases should be verified and certified (which may not be easy). These aspects were highlighted in the talk from KIST and their implementation of an Open Data environment but touched on by several speakers in passing. There is a big difference between useful information and useless information overload. The Big Data community has been very keen in defining a number of “Vs” which the data should comply with, like (i) Volume, (ii) Value, (iii) Veracity, (iv) Visualisation, (v) Variety, (vi) Velocity, (vii) Viscosity and (viii) Virality. These are for the quality and practical use of data. Some of these “Vs” could be easily adapted to materials science by giving more exact quantitative criteria. We recommend that in parallel with putting more efforts in ML and data-based modelling in materials science there should standards and agreements on data and methods to test and certify data from different sources.

2. Report of main activities

Session 1 - Data and/or physics I?

Session chair: Andreas Mauracher

Jens Berg (Uppsala University)

"Introduction to machine learning + Dara driven discovery of governing equations"

Particularly two complementary techniques within machine learning are found very powerful: Bayesian statistics and neural networks. Machine learning has become very easy and in principle you don't even need any programming experience.

Data sets can be approximated using linear or non-linear models (functions) where the initially unknown coefficients and weight factors (and other types of biases) are to be determined. By rotating the linear model 90 degrees gives a model called perceptron, an artificial neuron network, which can be of a single or a multilayer type and becomes finally an artificial neural network which can be used for example in supervised learning. The number of coefficients in neural networks grows very fast (in fact exponentially).

Training a network is a pure optimization process easily with billions of coefficients or so where the minimization is automatically done typically with gradient techniques. The user, however, has to decide initially the number of layers, and number of neurons in each layer as well as the learning rate, cost functions etc., which are called hyper parameters. Too many parameters cause easily a problem called overfitting, whereas too few parameters results in under fitting. Validation can be done by monitoring cost/loss functions for different models in minimizing.

Machine learning is big today. There is a large community out there and much open source software is available for CPU and GPU computers. The advice is to choose carefully among the ones with a long life expectancy.

David vad der Spoel (Uppsala University)

"Data + Physics = Models"

DATA: The speaker started by introducing the work behind his "virtualchemistry.org" initiative and presents a library, quantum chemical data base of molecular properties for force field development. Although not extensive in size yet it contains polarizabilities and other properties needed for molecular models used in simulations.

PHYSICS: Starts from Born-Oppenheimer approximation which is the basics of all molecular mechanical force fields and classical simulations where the electrons are treated with various "classical" models. Within BO the Hamiltonian can be formulated where the potential energy terms are associated to the force fields and the particles (atoms, ions, molecules etc) are moved based on Newtonian dynamics. Among the most intrigued interactions are the Coulombic due to distributed charges and multipoles replacing the quantum mechanical electron densities but also adding means for polarizabilities to mimic the redistribution of electrons due to various short and long-ranged interactions which can be done using different schemes.

The theme DATA+PHYSICS in this project applies while optimizing the all models and parameters. Machine learning is used in the process by applying standard techniques like regression and fitting but also Bayesian Monte Carlo. There were several take-home-messages such as: "Machine-learning without Physics is unlikely to work"

Session 2 - Insight from electrons

Session Chair: Olle Eriksson (Uppsala U)

Elsebeth Schröder (Chalmers Univ of Technology)

"Fundamental insights from non-local correlation of electrons"

The speaker introduces "the parameter free non-local vdW-DF method" as a general-purpose materials tool developed in her group focusing on non-local correlations and directional bonding, related to kinetic energy, and compares it with standard DFT to demonstrate its superiority. The features in the vdW-DF include consistent exchange vdW-DF-cx and spin in vdW-DF as well as the hybrid functional with some Fock exchange locally. To create a method with the correct "PHYSICS" has been important. The method is demonstrated by binding energies in noble gas dimer calculations and those in DNA base stacking. Kinetic signatures of the vdW binding were explained. The talk then continued with benchmarks of functionals to find the "Best of test".

Silvia Gallego (CSIC, Madrid)
"Towards a multiscale model of magneticexchange: the case of spinel ferrites"

The speaker introduces the concept of micromagnetics by building a matrix of mesh elements of both soft and hard types (20 times more hard) from underlying atomistic ab initio (DFT) calculations of electronic structures of the magnetic ground state. Spin Hamiltonian is created based on the atomistic models and applied on some cubic ferrite structures. Choice of spin-dependent XC functional is important in calculations of the electron gaps. The critical part is to map the DFT energy onto a Heisenberg model and include magnetic exchange interactions within a certain cut-off distance. The physics behind the magnetic exchange in focus as well as the multi-scale approach

Session 3 - Coupling & linking ("Multiscale modelling")

Session Chair: Alexander "Sasha" Lyubartsev (Stockholm U)

Vladimir Lobaskin (University College Dublin)
"Multiscale modelling of BioNanointerfaces"

The speaker's approach is: Start with bottom-up multiscale modelling and use the hierarchy of biomolecules and finally use data and statistical models. In describing the data-driven part he stresses the importance of deriving reliable descriptors from physicochemical properties of all inter-particle interactions and to find the statistical tools to establish the relationships between the properties. *Model reduction methodology should be developed to describe all types of NP-biomolecule interactions. Also QSAR types of approaches will be useful.*

Georg Schmitz (Access e.V., Aachen)
"Interactions between discrete models (electronic/atomistic/mesoscopic) and continuum models of microstructures"

Clearly no single simulation technique can handle the whole process. Multiscale and integrated modelling is needed. This requires largely standards in methods and communications and open simulation platforms. This is the very task of the EMMC. The speaker introduces then MODA (metadata scheme) and its further development the EMMO ontology.

Much of thermodynamic and concentration data is collected and used in simulations. Simulations with discrete models are used to create microstructure input to be used in continuum type of simulations. Physics + Models = Data. Since modelling is done as a chain of different methods and models it is important to estimate how errors and uncertainties may propagate in the process.

Daniel Höche (Helmholtz-Zentrum Geesthacht)
Multiscale modelling of corrosion in reinforced concrete structures"

Currently: The entire multiscale engineering model is developed (SINTEF). * Prediction of service life time is possible. * The Chemistry around it is to be implemented. * Experiments to calibrate the model are required * The Physics of self-healing of concrete structures is been developed.

Future: More thermodynamical data required- *Increased exchange of data between different models * Modelling of self-healing agents should be incorporated in the model * Integrate modelling with already existing commercial cements * Add

more environmental scenarios * Stable FEM code to simulate real-life infrastructure is being developed * Lab-scale testing to be accelerated.

Vladimir Rybkin (University of Zürich)
"Density Functional Embedding Theory: a Gaussian and Plane-waves Implementation for the Condensed Phase"

As most systems of interest are too large to do accurate quantum calculations other types quantum calculation schemes are needed. The speaker (working closely with Jurg Hutter) starts with an introduction of embedded approaches such as QM/MM, ONIOM, conquer and divide several other in studies of condensed phased systems using DFT. In particular those from the group of Emily Carter. He explains then the Wa-Yang functional to extend DFT to covalently bonded systems (from Carter's group) and outlines the formalism and algorithm and embedded potentials. Finally he shows the timing results from optimizing the scheme at different levels of theory.

The implementation will be soon available in CP2K. Also multiconfigurational multireference methods for bond-breaking will soon appear Linear response molecular properties will become available

Session 4 - Data and/or physics II?

Session Chair: Philippe Bopp (VISTEC)

Luca Ghiringhelli (Fritz-Haber Institute, Berlin)
"Novel materials discovery: big-data analytics methods and infrastructure for building maps of materials"

The author first presents "SISSO" (Sure Independence Screening and Sparsifying Operator), a compressed-sensing method for identifying the best low dimensional descriptor in a very large ab initio data showing some quantitative predictions of the ground-state enthalpies of octet binary materials and applying it to predict the metal/insulator classification of binaries in relation with experimental data. For example it rediscovers the available pressure-induced insulator-to-metal transitions and thereby allows the prediction of yet unknown transition candidates by experimental validation. The author then presents neural network prediction of nano islands and crystal structures. The author stresses the importance of physical information but more the great synergies of the combination.

Michele Ceriotti (EPFL, Lausanne)
"Machine learning like a physicist"

He visions ML as a sophisticated universal interpolator between a few known values of the properties. But can it be made accurate and self-consistent? He applies kernel principal component analysis (KPCA) to measure similarities between structures by obtaining a general-purpose similarity kernel. Research focus on fluctuations in materials and anharmonic free energies. Multiple impressive examples of benefits of clever use of ML methods. Story on what we can learn from the ML results.

Session 5 - Industrial perspectives

Session Chair: Jérôme Bikard (Solvay R&I)

Bart Rijksen (Dow Chemicals)
"An industrial perspective on materials modelling: physics-based versus datadriven"

What modelling tools (specialists using them) are used depends simply on (i) the cost (ii) speed (iii) availability. This Dow group does not do coding. In a typical toolbox there are: ML, First-principles, PDEs, Statistics, Simulations (atomistic,

continuum, finite element). The speaker shows some robust examples of modelling found very useful to understand complex structures through physics.

As take-away-messages: theory development not a primary goal and the choice of modelling tools/techniques is pragmatic. First-principles modelling provides better understanding. The combination is a double-win.

Manuel Sparta (Teknova AS, NORCE)
"Physics-based models within industrial metallurgy and opportunities for data driven modelling"

Problems are generally time limited. FEM is current workhorse. Used silicon-Mn pilot furnace as example. Many physico-chemical properties are strongly linked and there are large temperature differences to contend with. Digitisation and data driven modelling gives incomplete understanding. Combined theory and data driven models is better. Industry wants to use modelling but speed is of the essence.

James Goddin (Granta Design)
"Materials Digitization for BusinessDecision-Making"

Granta Design is a spin-off from Cambridge University offering a very broad portfolio of products and a network worldwide customers and collaborators within a niche called integrated computational materials engineering (ICME). They provide educational packages in intelligent materials design. This consists of Research-design/analysis-production-service (end of life).

Fredrik Hedman (Noruna AB)
"Computational Science needs Lean and Agile"

The speaker presents the so called waterfall model, which is used by many companies. However this is very inefficient multiple projects running in parallel. The project teams are often heterogeneous. Typically every project takes long time to finish and the results are only obtained at the very end. In this talk the speaker introduces the approach which produces "results" at all the time and become the final end-product.

Session 6 - Data and/or physics III?

Session Chair: Natalia Konchakova (HZG)

Seungchul Kim (Korean Inst. Sci. & Tech., Seoul)
"KIST's dream of an R&D informatics platform"

The KIST platform is aims to facilitate data-driven research for the development of new materials. A key issue here is to collect data and organize it as an easily accessible data structure. This would require research to be done on past research. He then explains the recent governmental plan for R&D management. This includes collecting and managing data from industrial PCs and servers after motivating researchers to share their data. However, it is not clear yet (to anyone) how to generate novel knowledge from archived large-scale data.

Session 7 - Towards force-fields for complex systems

Session Chair: Byung-Hyun Kim (KIER)

Martin Horsch (American U of Iraq, Sulaimani)
"Multi-criteria optimization for user adjustable molecular force fields"

He discusses massively parallel MD and scaling for trillions of particles. In physics based models the intermolecular interactions (short-range repulsion and dispersion and long-range electrostatics) of pair potentials are realistically

represented. In datadriven models with free parameters are used adjusted to thermodynamic data giving reliable prediction of fluid properties via interpolation and extrapolation. He then comments (criticizes) on formulations in the ROMM VI and MODA about fitting parameters in one and not the other of the two types of models. He finishes by saying the there is a paradigm shift in molecular modelling *from* where only a limited number of experts were developing force fields for different systems.

Dirk Reith (Fraunhofer Inst for Algorithms and Sci Comp, St Augustin, DE)
"Automated Global and Local Optimization Methods for Atomistic Force-Field Development"

As one author of iterative Boltzmann inversion he presents the method to compute effective potentials from RDFs where chemical intuition plays a big role. He then presents ESPResSo package to perform the CG simulations. Then he comes to Wolf2Pack FF optimization package also developed at Frauenhofer which is a very modular software and shows application of it on a very large database which gives the connection to data-based modelling.

Session 8 - ...and the future

Session Chair: Andreas Röckert (Uppsala U)

Itai Panas Chalmers (Univ of Technology)
"In anticipation of the unknown"

"The speaker discusses –among many other issues- that from the beginning we learn by playing (= trying) and the process continues all our lives. It also applies on modelling but can we divide it to data-driven, physics-diven, reality-diven or something else? Maybe it is EMMC-driven as we should formulate our project proposals in an endorsed fashion?!"

Report from the PhD students' panel discussion

Aatto Laaksonen Uppsala Univ/Stockholm Univ
"MODA+EMMO: EU-supported nomenclature to assist and systematize model development"

Peter Kusalik Univ of Calgary, Canada
"Big Data and Machine Learning: A Path to Physical Insight"

The speaker reminds us that there have been always so called "black box" approaches where user does not care what is going on in the box but the data coming out can be still very useful. ML can be seen very much as a more modern adaptive black box. As an example of self-adaptive algorithms he goes back to his own PhD project in 80's showing that he was already doing ML before the term was used very much. He chooses another example of how to make RDFs more accurate by adding the orientation of the molecules in a local coordinate system. This idea can be further generated to visualizing a reaction coordinates which he shows for a radical reaction. He then moves to the first example of nucleation of ice where the amount of data can be enormous and finding patterns in such big data is a non-trivial task. He shows finally that training the data needs to be well supervised and physical models will still be the key in informing and directing us in learning from data. Although ML is big today it only means that there are new tools and algorithms today. It is not really a new paradigm as started already 50-60 years ago.

3. General Discussion Feedback

The following topics were in focus.

- *Definition of terms*
- *Origin of data (data-driven models, data-informed physics models)*
- *Analysis of data-driven models with respect to materials modelling*
- *Definitions of machine learning in the field of materials science*
- Data-driven approaches etc. should be taught to physics/chemistry/materials science students

An excerpt from the discussion on the third topic is here. Overall the discussion was used to help formulating the recommendations from the meeting.

- How do you define the *data* ? For example Turing said that everything is data including the computer program". "Of course it is obvious that experimental results (spectroscopic etc) are data but how to define data more rigorously?
- In a generic sense we can start from binary digits zeroes and ones and these become numbers and words, values and structures etc. Data then obviously needs mathematics to form equations. We are talking about physics-based models like the equation for heat conduction etc. The conductivity data is then taken from some data base. And data bases themselves can contain simple values but also tables, complex structural and dynamical data.
- Then again we have physics-based with no coupling to any material specifically so it can be considered as a universal equation. But when you apply it on a specific material you need data to describe the structural and other molecule-specific parameters and potentials.
- What we do is simply constraining the data to become applicable for physics-based models. For example we have to be aware of not using data which violates energy and other conservation principles. At the same time we can apply initial and boundary conditions which we know in advance or direct our modelling towards specific correlations we are interested in in choosing appropriate data. In fact you can also check and verify that your data is correct for heat conduction equation.
- How do we see the Lennard-Jones potential? A physics-driven model? Yes, but the values to the potentials are the data-entry. Still can we call it a data-informed physics model as such. After all we don't want to change the name. Again, L-J potential is nothing but the materials relation and not a physics equation.
- Was it correct that the only physics models which are not data-informed are the ab initio methods as no empirical information by definition should go in?
- No, they are also data-informed. As you need to provide potentials and exchange correlation functionals. The potential which is the Coulomb potential, to be said is a universal ntity including the elementary charge but surely the exchange correlation functional is your choice, basis sets too.
- Today we have a large number of functionals making ab initio schemes (DFT) less universal.
- Don't forget that it is DFT not being universal but ab initio is. Meaning DFT is not ab initio as Pople always said. Although this is a different discussion we don't start here.
- So how does this work? Do the physical models generate the data to data-informed models?
- Now the data models search data sets for specific correlations. And if you ask a data model to search for a correlation like in heat conduction and the data points match in that kind of equation. That is we constrain the data model not to search just any kind of correlation but specifically that for heat conduction.
- So you are not examining the data set what types of correlation it contains in general but rather want to find a specific correlation?
- This depends on the constraints you apply in searching correlations in the data.

4. Programme

MONDAY, JUNE 11, 2018

(The Siegbahn hall at Ångström)

Color code: Keynote speaker | Invited talks and contributed talks

11.30–13:00	*Registration*	
13:00 – 13:20	*Welcome + purpose of the workshop (Kersti Hermansson)*	

Session 1: Data and/or physics? I Session Chair: Andreas Mauracher (Innsbruck U)

13:20 – 14:00	Jens Berg Uppsala University	*"Introduction to machine learning + Data-driven discovery of governing equations"*
14:00 – 14:10	Johan Tysk Vice-rector of S&T @ Uppsala Univ	*Address from the Dean of Science and Technology*
14:10 – 14:40	David van der Spoel Uppsala University	*"Physics + Data = Models"*

Session 2: Insight from electrons Session Chair: Olle Eriksson (Uppsala U)

14:40 – 15:10	Elsebeth Schröder Chalmers Univ of Technology	*"Fundamental insights from non-local correlation of electrons"*
15:10 – 15:30	Silvia Gallego CSIC, Madrid	*"Towards a multiscale model of magnetic exchange: the case of spinel ferrites"*
15:30 – 15:50	*Coffee*	

Session 3: Coupling & linking ("Multiscale modelling")

Session Chair: Alexander "Sasha" Lyubartsev (Stockholm U)

15:55 – 16:40	Vladimir Lobaskin University College Dublin	*"Multiscale modelling of BioNano interfaces"*
16:40 – 17:00	Georg Schmitz Access e.V., Aachen	*"Interactions between discrete models (electronic/atomistic/mesoscopic) and continuum models of microstructures"*
17:00 – 17:20	Daniel Höche Helmholtz-Zentrum Geesthacht	*"Multiscale modelling of corrosion in reinforced concrete structures"*
17:20 – 17:40	Vladimir Rybkin University of Zürich	*"Density Functional Embedding Theory: a Gaussian and Plane-waves Implementation for the Condensed Phase"*
17:40 – 17:55	*Further impressions from Sessions 1-3 (led by today's chairmen)*	
18:00 – 19:00	*Poster micro presentations*	
19:00 – 20:00	*Poster session and mingle*	
20:00	***Buffet dinner at Ångström*** *with some reflective comments from Magnus Ullner, LU*	

TUESDAY, JUNE 12, 2018

(The Siegbahn hall at Ångström)

Color code	Keynote speaker	Invited talks and contributed talks

Session 4: Data and/or physics? II Session Chair: Philippe Bopp (VISTEC)

09:00 – 09:45	Luca Ghiringhelli Fritz-Haber Institute, Berlin	*"Novel materials discovery: big-data-analytics methods and infrastructure for building maps of materials"*
09:45 – 10:30	Michele Ceriotti EPFL, Lausanne	*"Machine learning like a physicist"*
10:30 – 10:50	*Coffee (the posters are available)*	

Session 5: Industrial perspectives Session Chair: Jérôme Bikard (Solvay R&I)

10:50 – 11:10	Bart Rijksen Dow Chemicals	*"An industrial perspective on materials modeling: physics-based versus data-driven"*
11:10 – 11:30	Manuel Sparta Teknova AS, NORCE	*"Physics-based models within industrial metallurgy and opportunities for data-driven modelling"*
11:30 – 11:50	James Goddin GrantaDesign	*"Materials Digitization for Business Decision-Making"*
11:50 – 12:10	Fredrik Hedman Noruna AB	*"Computational Science needs Lean and Agile"*
12:10 – 13:40	***Lunch at Eklundshof*** *(Eklundshovsvägen 7) (official photo before leaving)*	

Session 6: Data and/or physics? III Session Chair: Natalia Konchakova (HZG)

13:40 – 14:25	Michael Probst University of Innsbruck	*"Accuracy vs. insight"*
14:25 – 15:10	Seungchul Kim Korean Inst. Sci. & Tech., Seoul	*"KIST's dream of an R&D informatics platform"*
15:10 – 15:40	*Further impressions from Sessions 4-6 (led by today's chairmen)*	
15:40 – 16:00	*Coffee (the posters are available)*	

Panel discussion

16:00 – 17:30	*Panel discussion:* ***Physics-driven vs. data-driven***

19:00	***Dinner at Norrland's Nation*** *(Västra Ågatan 14, i.e downtown)* *With the announcement of the poster prize*

WEDNESDAY, JUNE 13, 2018

(The Siegbahn hall at Ångström)

Color code	Keynote speaker	Invited talks and contributed talks

Session 7: Towards force-fields for complex systems

Session Chair: Byung-Hyun Kim (KIER)

09:00 – 09:20	Martin Horsch American U of Iraq, Sulaimani	*"Multi-criteria optimization for user-adjustable molecular force fields"*
09:20 – 09:40	Dirk Reith Fraunhofer Inst for Algorithms and Sci Comp, St Augustin, DE	*"Automated Global and Local Optimization Methods for Atomistic Force-Field Development"*

Session 8: … and the future

Session Chair: Andreas Röckert (Uppsala U)

09:40 – 10:10	Itai Panas Chalmers Univ of Technology	*"In anticipation of the unknown"*

10:10 – 10:50	*Coffee* + PhD students' panel discussion (open to all). The Siegbahn hall *in parallel with* *Coffee* + The Annual meeting of Theoretical Chemistry Section of the Swedish Chemical Society *(Peter Ahlström).* Room 4001
10:50 – 11:10	*Report from the PhD students' panel discussion*

Session Chair: Peter Broqvist (Uppsala U)

11:10 – 11:50	Aatto Laaksonen Uppsala Univ/Stockholm Univ	*"MODA+EMMO: EU-supported nomenclature to assist and systematize model development"*
11:50 – 12:10	Peter Kusalik Univ of Calgary, Canada	*"Big Data and Machine Learning: A Path to Physical Insight"*

12:10 – 12:25	*Further impressions from Sessions 7 and 8 (led by today's chairmen)*
12:25 – 12:40	*Conclusions + Closing of the workshop*
12:40 – 13:15	***Lunch sandwiches***
13:15 – 15:00	Open EMMC session for preparation of recommendations and report for the European Commission

The outcome of this workshop was compiled by	Kersti Hermansson (Uppsala University)

Workshop organisers	UU
Contributing beneficiaries	MDS, POLITO, UOY

EC-Grant Agreement	723867
Project acronym	EMMC-CSA
Project title	European Materials Modelling Council - Network to capitalize on strong European position in materials modelling and to allow industry to reap the benefits
Instrument	CSA
Programme	HORIZON 2020
Client	European Commission
Start date of project	01 September 2016
Duration	36 months

Consortium		
TU WIEN	Technische Universität Wien	Austria
FRAUNHOFER	Fraunhofer Gesellschaft	Germany
GCL	Goldbeck Consulting Limited	United Kingdom
POLITO	Politecnico di Torino	Italy
UU	Uppsala Universitet	Sweden
DOW	Dow Benelux B.V.	Netherlands
EPFL	Ecole Polytechnique Federale de Lausanne	Switzerland
DPI	Dutch Polymer Institute	Netherlands
SINTEF	SINTEF AS	Norway
ACCESS e.V.	ACCESS e.V.	Germany
HZG	Helmholtz-Zentrum Geesthacht Zentrum für Material- und Küstenforschung GMBH	Germany
MDS	Materials Design S.A.R.L	France
QW	QuantumWise A/S	Denmark
GRANTA	Granta Design LTD	United Kingdom
UOY	University of York	United Kingdom
SYNOPSYS	Synopsys Denmark ApS	Denmark

Coordinator – Administrative information	
Project coordinator name	Nadja ADAMOVIC
Project coordinator organization name	TU WIEN
Address	TU WIEN | E366 ISAS | Gusshausstr. 27-29 | 1040 Vienna | Austria
Phone Numbers	+43 (0)699-1-923-4300
Email	nadja.adamovic@tuwien.ac.at
Project web-sites & other access points	https://emmc.info/