Institution: Durham University		
Unit of Assessment: UoA 9: Physics		
Title of case study: CASTEP: A materials modelling code with wide ranging industrial and academic applications and commercial success		
Period when the underpinning research was undertaken: Between January 2000 and December 2020		
Details of staff conducting the underpinning research from the submitting unit:		
Name(s): Prof. Stewart Clark	**Role(s) (e.g. job title):** Professor of Physics	**Period(s) employed by submitting HEI:** Oct 1997 to present
Period when the claimed impact occurred: Between August 2013 and December 2020		
Is this case study continued from a case study submitted in 2014? Yes		

1. Summary of the impact

Research in Physics at Durham, led by Prof. Stewart Clark, focused on the development of computation methods to calculate the electronic, physical and chemical properties of materials from first principles. These are incorporated in the CASTEP code, whose on-going development has worldwide impact beyond academia. Professor Clark is one of the six original and current co-developers of CASTEP, which can be used in a variety of fields, from understanding semiconductor devices and light emitting displays, to areas as diverse as the behaviour of Earth minerals, catalysis and drug design. CASTEP was commercialised for use under license by Biovia Inc. and has been purchased by approximately 1,000 high-tech companies developing new materials across chemical, pharmaceutical, auto and jet engine industries. Total sales revenue for Biovia (Dassault Systems) resulting from direct, industrial impact of CASTEP is in excess of USD3.25million per year in the current REF period.

2. Underpinning research

The properties of materials can be predicted from first principles (quantum mechanics) from their electronic structure. However, it is prohibitively time consuming to calculate these from a wavefunction approach for multi-electron systems. Instead, density functional theory (DFT) is a much more powerful technique to solve many body problems in quantum mechanics. The CASTEP research tool uses DFT to calculate the electronic properties of materials and provides information that can be compared with directly measurable data, such as how the material responds e.g. to light, phonons, neutrons etc. Critically, in this area of research the border between academia and industry is porous with fundamental research having immediate impact and take-up in industry via CASTEP.

The background to this case study is that the CASTEP computer code was started in the late 1990s by Prof. Stewart Clark (a member of Durham Physics Department from 1997 to present) and a small group of theoretical/computational condensed matter physicists. They realised that the lack of efficient computational tools with a modern code design was a serious barrier to their research, and developed a new code using advanced software engineering techniques to provide a robust, efficient, and portable code designed to run on high performance parallel computers. Their commitment to full documentation also allows additional functionality to be easily incorporated by external researchers as well as the authors. The original CASTEP Developers Group consisted of Stewart Clark (Durham), Chris Pickard (Cambridge), Phil Hasnip and Matt Probert (York) and Matt Segall (now in industry), joined later by Keith Refson (Royal Holloway). While all are credited as equal co-authors, Prof. Clark's contribution is the on-going development of new, efficient computational implementation of exchange-correlation functionals and of density functional perturbation theory, which are the accurate approximations underlying the techniques for property prediction that form the heart of the CASTEP code. He remains responsible for the band structure and exchange-correlation functional coding, is the developer of the cell and Fast

Fourier Transform (FFT) modules and a co-author of the linear response code (such as phonons and thermodynamics). The first release (academic and industrial) of CASTEP was in 2001, with the paper describing its properties published in 2002 [R1].

Since 2001, Prof. Clark and the other developers have applied CASTEP to a wide range of fundamental research questions, and continually update and upgrade the code, with the underpinning research activities published in the most relevant international peer reviewed journals and the scientific significance of this research widely recognised, e.g. [R2] which is cited over 8,300 times.

Recent research incorporated into CASTEP by Prof. Clark since 2014 includes work on highly correlated electron systems. DFT relies on approximate exchange-correlation functionals, which reduces its ability to reproduce strongly correlated electron physics in many materials. The exact exchange potential should instead be used, but gives a first order correction term which is a slowly convergent infinite sum, so solutions are plagued with numerical instabilities. Instead, Clark developed a new, much more computationally efficient way to solve non-local interactions between electrons [R3]. This improves the accuracy of the band structure calculations, particularly of metal-semiconductor interfaces, impacting on new device technologies. However, even this approach is not effective in elements with open f-shells (rare earths) where the electrons are spatially concentrated. These are important e.g. in advanced magnetic materials for data storage. Clark showed that using Dynamical Mean-Field Theory rather than DFT alone allowed accurate calculations of the structures of these materials [R4]. Both these new techniques are now incorporated into CASTEP.

Another set of research led to improved functionality for terahertz (THz) time-domain spectroscopy in CASTEP. THz radiation is well tuned to semiconductor band structures, so is an important probe of these materials. It is also well matched to vibrational modes in DNA and other organic molecules yet passes easily through packaging/clothes and is non-ionising, leading to its use in the new generation of airport body scanners. Yet there are relatively few materials which can generate useful THz radiation, and these are generally expensive and/or difficult to produce. Instead, Silicon Carbide is cheap, and can be easily manufactured into lattices. CASTEP calculations by Clark showed that lattice vibrations (phonons) in these non-linear crystals could give rise to THz radiation, as confirmed by experimentalists [R5].

Prof. Clark and the Group have paid particular attention to the accuracy of their calculations showing that codes such as CASTEP based on recent DFT methods yield reproducible results [R6]. Older DFT implementations and other DFT codes predicted different values, even codes which ostensibly solved the same equations. The recent improved precision and reliability makes the CASTEP code a high value resource, knowing results are reliable, enhancing communication and collaboration between theorists and experimentalists, in both industry and academia.

The continual development of CASTEP has increased its applications in research. The code has been used in over 17,000 peer-reviewed publications (Google Scholar, August 2020) across a wide range of disciplines. It is also used in training highly-skilled graduate students, many of whom have gone on to work in industry. Several hundred PhD theses from 2002 to the present have been based on calculations using CASTEP, contributing to the large number of scientific papers published in leading peer-reviewed international journals.

3. References to the research

[R1] First-principles simulation: ideas, illustrations and the CASTEP code, MD Segall, PJD Lindan, MJ Probert, CJ Pickard, PJ Hasnip, SJ Clark, MC Payne, Journal of Physics: Condensed Matter 14. 2717, 2002 https://doi.org/10.1088/0953-8984/14/11/301 Over 9300 citations.
[R2] First principles methods using CASTEP, SJ Clark, MD Segall, CJ Pickard, PJ Hasnip, MIJ Probert, K Refson, MC Payne, Zeitschrift für Kristallographie 220, 567, 2005. https://doi.org/10.1524/zkri.220.5.567.65075 Over 8300 citations.

[R3] Self-interaction free local exchange potentials for metallic systems, SJ Clark, TW Hollins, K Refson and NI Gidopoulos, J. Phys. Condens. Matter 29, 374002, 2017. https://doi.org/10.1088/1361-648X/aa7ba6
[R4] Many-body renormalisation of forces in f-electron materials, E Plekhanov, P Hasnip, V Sacksteder, M Probert, SJ Clark, K Refson and C Weber, Phys. Rev, B 98, 075129, 2018. https://doi.org/10.1103/PhysRevB.98.075129
[R5] Terahertz time-domain spectroscopy of folded acoustic phonons in 4H and 6H silicon carbide, AT Tarekegne, B Zhu K Kaltenecker, K Iwaszczuk, SJ Clark and PU Jepsen, Optics Express 27, 3618, 2019. https://doi.org/10.1364/OE.27.003618
[R6] Reproducibility in density functional theory calculations in solids, K. Lejaeghere, SJ Clark, et al, Science 351, aad3000, 2016. https://doi.org/10.1126/science.aad3000.

The code has been used in over 17,000 peer-reviewed publications (Google Scholar, August 2020) across a wide range of disciplines in leading peer-reviewed international journals. The highly cited nature of these referenced publications provide evidence for the quality of the research.

4. Details of the impact

The impact of CASTEP [E1] comes from its ability to accurately simulate a wide range of materials including crystalline solids, surfaces, molecules, liquids and amorphous materials. It is a crucial piece of scientific equipment to researchers in materials science, in industry and academia, and analogous to key experimental tools and facilities in R&D. It can calculate the properties of any material that can be thought of as an assembly of nuclei and electrons with the only limitation being the finite speed and memory of computer systems. Examples of impact areas where CASTEP is used are:

- Device technologies: modelling materials for new semiconductor devices including nanoparticles, high dielectric constant materials, data-storage materials
- Geology: elastic, thermal and dynamical properties of Earth and planetary material under extreme conditions
- Spectroscopy: interpretation of experimental studies in a wide variety of spectroscopic techniques such as Raman, infra-red, nuclear magnetic resonance, electron energy loss spectroscopy, neutron spectroscopy
- New light emitting materials: investigations on liquid crystals, light emitting polymers, light emitting semiconducting materials all used in new display technologies
- Structure prediction: new materials ranging from pharmaceuticals and life-science molecules to semiconductors, spintronics and complex metals to name a few.

The reputation of the code quickly spread and the authors were approached to licence the code in 2001 by Accelrys Inc., a NASDAQ-quoted research and development software company based in San Diego, USA. Accelrys marketed CASTEP for sale in its Materials Studio modelling software portfolio for commercial scientific, industrial and technological applications as well as supplying a free licence for academic use of CASTEP. In 2014, the French maker of computer-assisted design programs, Dassault Systèmes, acquired Accelrys, whose 2013 revenue was USD122million, in a cash sale for USD750million. Now called the Biovia Materials Studio modelling and simulation platform, this has been bought by over 500 companies worldwide [E2, E9] since 2015, making this the best-selling software package of its type within industry. Purchasers include some of the largest manufacturers in the chemical (e.g. Johnson-Matthey), pharmaceutical (e.g. Pfizer), automobile (e.g. Toyota, General Motors) and user technologies (e.g. Sony) sectors, as well as smaller scale high-tech science companies, testifying to the commercial viability of the product across a wide spectrum of business sizes and sectors. On Prof. Clark's contribution, BIOVIA state: *"…rich functionality and reliable performance of CASTEP make it a popular application among academic and industrial users of BIOVIA software. Your personal contributions address some of the key areas that make this package attractive to customers of Dassault Systèmes BIOVIA."* [E2]

Research into the effectiveness of CASTEP, sponsored by Dassault Systèmes (Biovia), found that customers were able to recoup their investment in software tools up to 10 times over. The major

cost savings in applying computer simulations to the research and development of materials were from circumventing the need for costly experiments and led to shorter developmental timescales [E3, E4].

Sales of CASTEP are producing an annual return of USD3.25million (2014 to 2019) to Biovia, with total sales of CASTEP for commercial use now in excess of USD40million [E2]. This is a far from insignificant fraction of Biovia's annual turnover, which analysts currently estimate to be USD155million, with over 700 employees [E5]. However, with the code so deeply embedded in industry, it is impossible to quantify its full financial impact, since much of the information concerning R&D gains from use of CASTEP is commercially sensitive and not public. It is therefore necessary to identify proxies for the commercial impact. Enquiry of the Espacenet database of the European Patent Office reveals that since August 2013, 134 patent filings have been published which specifically refer to CASTEP and 945 on Google Patents [E6]. They are from a substantial range of countries, and range across a wide variety of materials. Examples include advances in semiconductor innovations, the manufacture of oxide films and the manufacture of liquid crystal display devices. CASTEP was specifically highlighted due to its impact in the most recent EPSRC Delivery Plan [E7].

A proxy for the significance of CASTEP's impact and reach is the huge size of the CASTEP reference database that is available on the Biovia website [E8]. Since 2013, there are over 500 papers relating to geological science and over 900 concerning drug design, many co-authored by industrial scientists, using Prof. Clark's recently added functionality. Free academic licencing for academic users is now obtained via a licence for use directly from Cambridge Enterprises Ltd. The fact that the academic licence is free and the intellectual property licenced is virtual, enhances the value of CASTEP as a research tool, which is seen in its monetary value when licensed for commercial use. In this respect Cambridge Enterprises Ltd. is wholly subsidising the commercial sale of CASTEP to academic researchers.

5. Sources to corroborate the impact

[E1] The CASTEP website (www.castep.org).
[E2] Letter supplied by Dassault Systèmes (dated 27 November 2019).
[E3] BIOVIA Materials Studio website (https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/biovia-materials-studio/, retrieved 24 September 2020).
[E4] European Materials Modelling Council (EMMC) Case Study: First-principle simulations of electronic structure in semi crystalline polyethylene (https://emmc.info/wp-content/uploads/2018/12/EMMC_ABB_CASESTUDY.pdf).
[E5] Wikipedia webpage for BIOVIA (https://en.wikipedia.org/wiki/BIOVIA, retrieved 24 September 2020)
[E6] Search results for 'CASTEP' from European Patent Office (24 September 2020) https://worldwide.espacenet.com/patent/search?q=CASTEP
[E7] EPSRC Delivery Plan 2019, page 7 (https://epsrc.ukri.org/newsevents/pubs/deliveryplan2019/).
[E8] https://www.materials-studio.com/products/collaborative-science/biovia-materials-studio/references/castep-references/
[E9] Letter of support supplied by BIOVIA (dated 30 September 2020)