Climate Adaptation Digital Twin HPC+QC Workflow

Abstract

A Digital Twin is a virtual model of a given system or process, where one as accurately as possible attempts to reproduce its behaviour. The goal is to model the time evolution based on initial values and possible perturbations. Some of the most ambitious projects of this type are found within the Destination Earth initiative, where digital twins of the Earth are constructed. Modelling processes at this scale requires massive amount of computational resources. Only now, with the advent of pre-exascale and exascale supercomputers, has this become a realistic prospect. Here, we scrutinise the present state-of-the-art in climate modelling, the Climate Adaptation Digital Twin (ClimateDT), which is currently being prepared for running on the most powerful supercomputing infrastructures in the world. We discuss the potential of quantum computers for further improving both the efficiency and accuracy of the models, by combining high-performance computing (HPC) and quantum computing (QC) in a hybrid HPC+QC manner. The biggest promise of quantum acceleration comes from aiding machine learning models and from improving the climate models by including processes that presently are computationally intractable due to their complexity. These include, for example, chemical processes in the atmosphere.

7 Introduction

The Climate Adaptation Digital Twin (ClimateDT) is a project issued by the European Centre for Medium-Range Weather Forecasts (ECMWF) in the Destination Earth initiative, where the goal is to develop a highly accurate digital model of the Earth see Figure 7. The aim is to develop an accurate model of the Earth in order to monitor and simulate the interactions between the natural environment and human activities with as high precision as possible. Through this, the effects of various natural phenomena and human actions on the climate can be studied. The underlying goal is to move from plausibility assessments of local and regional climate to fully-developed risk assessments. The ClimateDT is being developed in response to the European Commission's Green Deal and Digital Strategy and it will make it possible to predict the effect of specific climate actions, which will aid policy makers to make informed decisions on how to best mitigate the effects of climate change.

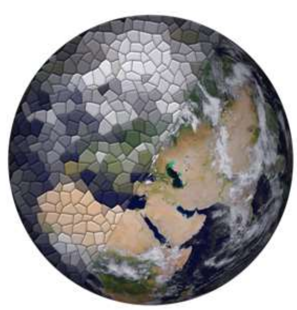

Figure 7: Digital twins of the Earth attempt to replicate the behaviour of certain aspects of the planet based on Earth Observation data and physical models.

The ClimateDT is being developed around two Earth System Models (ESMs), the Max Planck Institute for Meteorology's (MPI-M) Icosahedral Nonhydrostatic Weather and Climate Model (ICON) and ECMWF's Integrated Forecast System (IFS), the latter coupled with two different ocean models: FESOM and NEMO. Both of these ESMs have demonstrated the ability to run at grid scales finer than 5 km globally, coupled to an ocean model. The Climate DT introduces the idea of a generic state vector (GSV), which is evolved by the ESMs, quality controlled and interpolated to a common grid (5km or finer global mesh), and "streamed" to applications. This creates an information system that can scale across an unlimited number of applications that have access to all necessary data and achieve the long-sought goal of interactivity and new ways of co-design.

The ESMs in ClimateDT are being developed in three configuration: a coarse 10 km grid resolution for development, 5 km grid for production, and 2.5 km grid for prototyping subsequent Destination Earth phases. The ClimateDT focuses on five use cases (with a couple of keywords) drawn from climate impact sectors: forestry (wildfires, forest management), urban environments (heatwaves and heat island effect), hydrology (river flows, fresh water availability), hydro-meteorology (extreme-events, flooding), and energy (changing patterns of wind and sunshine, storm vulnerability). The ClimateDT is also complemented by other projects, such as the biodiversity digital twin BioDT [https://biodt.eu/].

Accurate digital twins of the entire Earth have only become possible with the latest generation of supercomputers, that is, the pre-exascale and exascale systems. There are presently five supercomputers on the Top500 list (June 2023) that have a sustained performance of over 100 petaFLOPS, that is, the capacity to perform over 10^{17} floating point operations per second (two of these, LUMI and Leonardo, are European). Even with this impressive increase in data processing capacity, an increase of a factor of million over the last 25 years, digital twins are an ambitious undertaking, and the models necessarily include approximations that affect the accuracy and reliability of the predictions.

Here, we scrutinise the present state-of-the-art workflows for setting up digital twins for climate adaptation, with the intent of identifying areas where quantum computing has a potential for speeding up or increasing the accuracy of selected parts of the entire simulation. Further, we identify the sources of the largest uncertainties in the model in the form of missing parameters or physics in the model, again with the aim to identify areas where quantum computing could provide an advantage.

8 Climate Digital Twin Workflow Analysis

8.1 Present Classical Approach

The Climate Digital Twin workflow is presented in Figure 8. The workflow begins with the typical initialisation and preparatory steps required by a climate or Earth System Model (ESM). In the Climate DT project, the ESMs in use are ICON and IFS. In the workflow, the current model state, illustrated as a Model State Vector (MSV), is propagated forward in time to produce a new state and, simultaneously, the model output or Output State Vector (OSV). This output is streamed (not saved) through a processing pipeline – that introduces additional diagnostic variables and handles interpolation, meta-data conversion and simple operations on the fields – to generate a Generic State Vector (GSV). The GSV is saved directly to Fields DataBase, which is a domain-specific object store developed at the ECMWF; another streaming approach is also being developed with the use of Maestro (https://www.maestro-data.eu/). The GSV is then forwarded to the applications and quality assessment and uncertainty quantification (AQUA), all of which can also utilise external data sources, e.g., observations, climatologies and reanalysis.

The most resource-heavy and time consuming part of this workflow, i.e., the bottleneck, is the climate model itself. Here we note, that the amount of data in a climate model is large. With a typical grid resolution of 10 km, the total number of grid points representing the atmosphere is in the hundreds of millions. Each grip

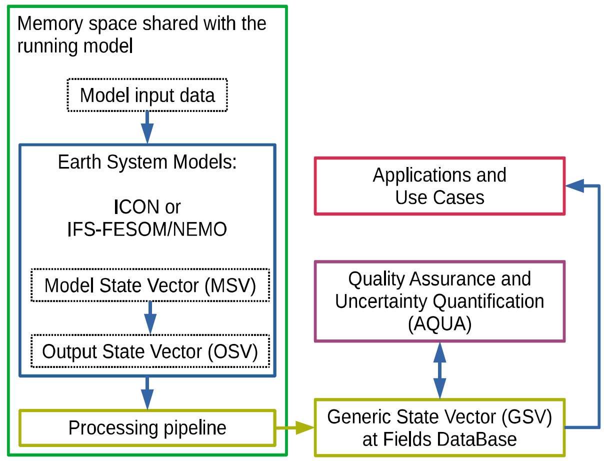

Figure 8: Operational version of the ClimateDT workflow that will be developed during 2023.

point has several variables associated with it, such as air density, temperature, wind speed, humidity, etc. The total parameter space is thus counted in the billions.

Figure 9 shows the relation between different processes in the ICON-Sapphire Earth system model Hohenegger et al. [2023]. What can be seen is that different processes are updated at different intervals, that is with different Δt. This is partly due to the varying computational complexity for propagating specific processes in time in the Earth and climate models. The shortest time steps are those of the dynamical core computations that solve the fluid dynamics equations of atmospheric motions, while the radiative transfer computations have the longest time steps. There is roughly a 1:30 ratio between the shortest and longest time steps. In the latest climate models within ClimateDT, with a resolution of 10 km, the time steps for dynamics and radiation are typically 60 s and 30 min, respectively. Presently, the wall-time for computing the individual time-steps range from the subsecond regime to around 10 s on the LUMI supercomputer. We note that doubling the resolution of the model typically requires halving of the time steps, following the Courant-Friedrichs-Lewy condition Courant et al. [1928]. Thus, doubling the resolution, e.g., going from 10 km to 5 km increases the computational complexity roughly by a factor of 8.

8.2 Quantum Perspective

From the previous section, we can identify two main challenges that hamper direct adoption of quantum computing to climate modelling problems within ClimateDT:

1. ”big data” problem
2. short wall-time for individual calculations

First, the climate models work on a large amount of data, both as input and output. These ”Big data” problems are, however, not directly suitable for quantum computers. The strength of quantum computers lies in being able to solve problems with a *moderate* amount of both input and output variables, where the relation between input and output variables is a highly complex equation that can be solved efficiently by some quantum algorithm, exploiting quantum parallelism Hoefler et al. [2023]. In other words, quantum computing typically requires problems that have a large potential solution space, but only a small set or even a single solution, with the additional provision that the input parameters need to be of the same order of magnitude as the number of qubits in the system.

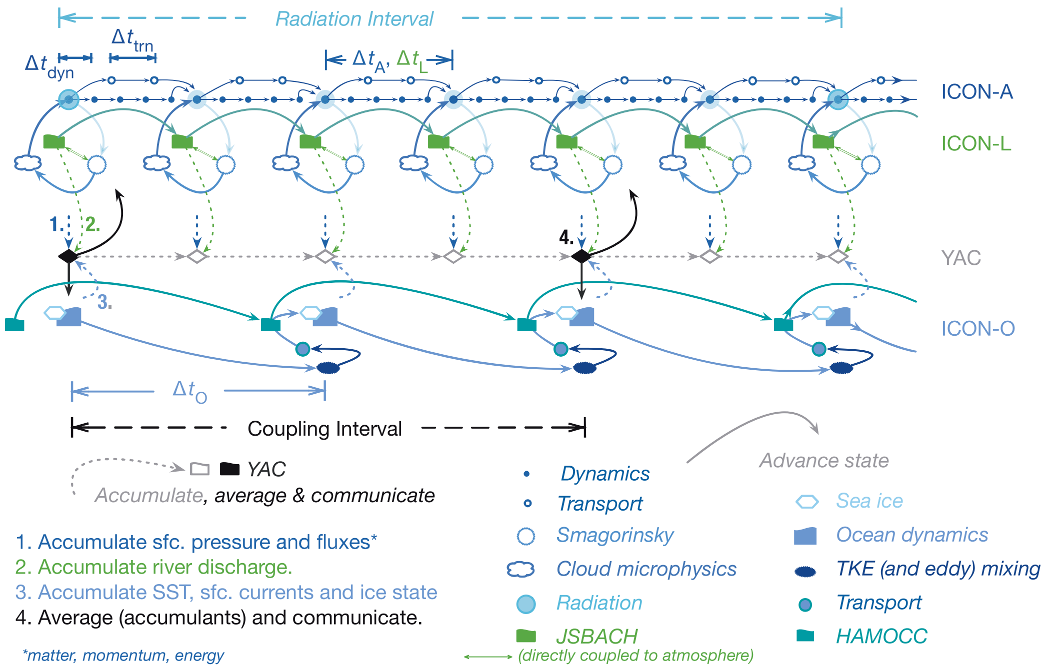

Figure 9: Time stepping in the ICON-Sapphire Earth system model. ICON-A: atmosphere component; ICON-L: land component; YAC: atmosphere-ocean coupler; ICON-O: ocean component. Reproduced from Hohenegger et al. [2023] under the Creative Commons Attribution 4.0 license.

Second, for quantum computers to be able to show a wall-time advantage over classical computers, they need to solve sufficiently complex algorithms. This means that the algorithms have to be sufficiently deep, that is, the number of basic operations has to be high. In practice, this means that single useful quantum computing calculations will take at least seconds to complete Humble et al. [2021]. Individual variational circuits can and do take shorter time, but the wall-time to solution is of course much longer, as several iterations need to be performed. On the other hand, already now, the shortest individual time-steps in the climate digital twins take less than a second, and even the longest around 10 seconds. Further, the aim of the ClimateDT initiative is to speed up the individual time steps significantly, with up to a factor of one hundred. This would push *all* of the individual propagation calculations into the sub-second regime. Thus, quantum computers cannot speed up these calculations further, as they already are faster than the fastest useful quantum computer calculations.

Climate models would thus, at a first glance, seem to be rather unsuitable for quantum acceleration. In order to gain some quantum advantage, we need to consider the problem at hand from a broader perspective. Simply taking present classical algorithms and the approximations they include and rely on, and transforming these to quantum versions of the same will not work. Instead, quantum advantage will be found by approaching the problem from different, new angles, utilising the unique features of quantum machines.

A large part of the calculations in the current workflows are in effect Computational Fluid Dynamics (CFD). Here, we have a direct connection to solving linear systems of equations. The HHL quantum algorithm for linear systems of equations, named after its authors Harrow, Hassidim, and Lloyd Harrow et al. [2009], and variations thereof, thus have the potential to speed up also CFD simulations. As noted by Lapworth Lapworth [2022], classical algorithms running on supercomputers are highly efficient at solving matrix equations by, for example, side-stepping the need for matrix inversions. Quantum algorithms do not need to, even *should* not rely on the same approximations as classical algorithms, however. Quantum algorithms like HHL and the Quantum Singular Value Transformation (QSVT) Gilyén et al. [2019] can efficiently perform direct matrix inversions, and should therefore be utilised for quantum advantage. The approach presented by Lapworth Lapworth [2022] relies on fault-tolerant quantum computers, but also hybrid classical/quantum algorithms for the NISQ era have been proposed and discussed Kyriienko et al. [2021].

8.3 Enhancing Machine Learning Approaches

The use of Machine Learning (ML) and Artificial Intelligence (AI) in climate modelling and related fields is a hot topic of research Tuia et al. [2023]. ML and AI show promise for accelerating the resource heavy calculations involved also in digital twins of the Earth Chantry et al. [2021]; Watson-Parris [2021]. Computational Fluid Dynamics (CFD), as discussed above, is central to climate models. A promising approach is, e.g., to decrease the grid size without losing accuracy, by using ML for interpolation Kochkov et al. [2021]. As the atmospheric events to a large part are rather smoothly changing, the speed-up from machine learning can be expected to be significant.

The connection to quantum computing here comes at a general level. As discussed in other sections of this report, quantum machine learning has potential advantages over purely classical machine learning. Also here, we need to keep in mind that present-day classical machine learning approaches are immensely powerful. Therefore, quantum algorithms will have a hard time directly competing with classical algorithms from a pure speed-up perspective Schuld and Killoran [2022]. Instead of speed-up, quantum computers can possibly increase the accuracy of the models or decrease the amount of training data required for building the models. Here, the advantage arises from doing the training *differently*, not necessarily faster. For example, hybrid quantum/classical neural networks, where a neural network consists of both classical and quantum layers, has the potential to outperform purely classical and purely quantum approaches Xia and Kais [2020]; Arthur and Date [2022]. As another example, Quantum Support Vector Machines (QSVM) have the potential to perform classification tasks more efficiently than their classical counterparts Havlíček et al. [2019]; Schuld and Killoran [2019]; Jäger and Krems [2023].

The understanding of where classical ML and AI can be utilised in climate modelling is thus being established through significant global efforts. The next step is to identify those machine learning tasks that can benefit from a quantum ingredient. This task is, however, highly empirical by nature. Only by testing, trial, and error, can the most successful quantum machine learning approaches be identified and refined. At this stage, it is too early to predict the future impact of QML on climate modelling. There is cause for careful optimism however, especially for improving the quality of the models, if not directly the time required for establishing them.

9 Missing Physics in the Models

In this section, we discuss two of the major approximations in the present climate DT models: clouds and atmospheric chemistry. Their inclusion is presently prohibitively expensive from a computational resource point-of-view. Quantum algorithms and quantum computing could bring about the necessary reduction in required computational resources in order to enable the inclusion of more parameters and additional physics into the climate digital twins, also beyond these two examples.

9.1 Clouds

Cloud feedbacks and cloud-aerosol interactions are the most likely contributors to the high values and increased range of equilibrium climate sensitivity in CMIP6 Meehl et al. [2020]. In the past, clouds have been poorly represented in Earth System Models (ESMs) due to the complex cloud formation process and because the models could not be run on the scales at which clouds form. Additionally, numerical cloud modelling has relied on the Eulerian continuous medium approach for all cloud thermodynamic variables. However, recently modelling has shifted towards Lagrangian particle-based probabilistic approaches in small and cloud-scale simulations. Clouds are being taken seriously – the World Climate Research Programme has launched a Grand Challenge on Clouds, Circulation and Climate Sensitivity and NASA has a Grand Challenge "Uncertainty Project" Fridlind et al. [2021] tackling cloud physics knowledge on ESMs.

Clouds are also a focus point for the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND) initiative, where a relatively recent review Stevens et al. [2019] proposed a protocol for the first intercomparison project of global storm-resolving models. The review presents 40-day global model simulations (these include ICON and IFS) with a grid resolution uniformly lower than 5km and addresses both scientific aspects and computational performance analysis. The outlook is optimistic even though the authors note that fully resolving shallow cloud systems, whose vertical (and hence horizontal) scale may be only a few kilometers, requires substantially smaller grid distances. This ties in with machine learning efforts for cloud cover modelling Grundner et al. [2022], and consequently with quantum machine learning efforts discussed above. We expect cloud representation to improve in all ESMs, including ICON and IFS; In the first phase, using purely classical supercomputing, and subsequently, by quantum-accelerated HPC.

9.2 Atmospheric Chemistry

Li et al. report that "climate models indicate at least a 30% uncertainty in aerosol direct forcing and 100% uncertainty in indirect forcing due to aerosol–cloud interactions" Li et al. [2020]. The accurate modelling of atmospheric chemistry would thus be of crucial importance for increasing the reliability of the climate digital twins. Atmospheric chemistry is highly challenging from a modelling perspective. Many of the reactions involve radicals, and several are photochemical in nature. This necessitates the use of highly sophisticated electron-correlation methods for describing the electronic structure of the molecular species. Highly correlated quantum chemical wave-function methods are notoriously difficult for classical computers. This is due to the expensive scaling of the so-called multi-reference methods that are required. These scale exponentially with the size of the problem (effectively, the number of electrons) at the limit of sufficient accuracy.

On the other hand, the electronic structure problem is naturally suited for quantum computers Cao et al. [2019]. Quantum chemical simulation is one of the major areas of research and development on all platforms of quantum computing, from quantum annealers, via quantum simulators, to general-purpose quantum computers. Quantum phase estimation (QPE) can provide sub-exponential solutions to electronic structure problems, but requires fault-tolerant quantum (FTQ) computers, due to massive circuit depth requirements. For NISQ devices, several iterative, variational algorithms have been devised, and progress is rapid. The importance of error-mitigation is recognised Glos et al. [2022]; Cai et al. [2023]; Kim et al. [2023], which gives hope for notable quantum advantage in the simulation of highly correlated electronic structure problems already before FTQ computing becomes a reality.

Having fast, highly accurate methods for simulating atmospheric chemistry is crucial, as the number of possible reaction pathways also grows rapidly with the size of the molecules involved in the reactions. With present-day supercomputers and modern classical quantum-chemical approximations, it is feasible to model a limited set of possible reaction pathways for smaller individual molecules. This is sufficient for demonstrating the importance of including chemical reactions and their interactions with the rest of the climate system in the models. In order to be of predictive accuracy and for decreasing the related uncertainties, massive improvements in modelling methods and capacity is, however, needed. Here, quantum algorithms and quantum computers can play a decisive role.

On-the-fly calculations of chemical reactions in the atmosphere within the climate models will remain out of reach for a long time, even with powerful quantum computers of the future. The run-time of individual quantum computing subroutines, including the necessary pre- and post-processing of data, will be significant, especially for cases where any quantum-advantage can be expected.

For climate modelling, where the average wall-time spent on each time step in the state evolution needs to be sufficiently short (below 1 second), atmospheric chemistry can be included in a parameterised manner instead. For creating accurate ML approaches for atmospheric chemistry, huge amounts of training data is required, and this can often only be obtained by performing quantum chemical calculations Kubečka et al. [2023]. Then, the advantage of quantum computers would come from enabling unprecedented accuracy for the data that forms the basis for the models.

10 Combining High-Performance Computing and Quantum Computing: HPC+QC

There are presently major ongoing efforts around the globe for connecting HPC infrastructure with quantum computers. This is perhaps even somewhat surprising, considering that quantum computers presently cannot solve any useful real-world modelling problem more efficiently than a single node of a supercomputer. At the same time, it is testament to the potential, and the *belief* in the potential of quantum computing for scientific modelling.

In Europe, the plans for making quantum computing relevant for research and development in academia and industry alike have been outlined, with the goal of having a European quantum computing infrastructure exhibiting quantum advantage by 2030. The first quantum simulators are already being integrated with HPC infrastructure in the HPCQS project [https://www.hpcqs.eu/]. In June 2023, the EuroHPC Joint Undertaking has signed hosting agreements for six different quantum computers to be placed in HPC centres around Europe, with the plan to make these available to European users in 2024. These first quantum computers are only the beginning, several updates and new procurements are already planned.

The actual (future) HPC infrastructure and its implementation needs to be accounted for. Already in the near-term, it is expected that individual supercomputers will be connected to several quantum machines of various types and implementations Johansson et al. [2021]. The initial setups, with individual QPUs distributed throughout the continent, connected to an HPC system, can be seen as precursors to a future where QPUs will be connected in parallel, either entangled or not. Plans for even tighter, on-chip integration of QPUs with classical processing units already exist, and may well be the way for reaching fault-tolerant quantum computing. With this in mind, more emphasis on developing parallel quantum algorithms, which simultaneously utilise several QPUs, in an HPC+nQC manner, would seem appropriate. For time-evolution problems like climate modelling, this can be a necessary development at a relatively early stage, in order to enable the quantum processing part to keep up with the classical computing tasks at each time step.

Reassuringly, the importance of investing in software development for hybrid HPC+QC applications has been recognised. These developments complement the efforts for developing purely classical software for exascale supercomputers and beyond, exemplified by the Destination Earth initiative.

Here, it is apt to note that there is a need for significant classical software development alongside the quantum algorithm research. Presently, pre- and post-processing tasks take up a significant portion of the total wall-time of executing a quantum algorithm. As an example, the recent experiment on spin dynamics using IBM's 127 qubit QPU, the actual time spent on the QPU was 5 minutes, while the wall-time of the experiment was a hundred times longer, over 9 hours Kim et al. [2023]. These overheads will decrease in the future, but at the same time, increasing qubit count will again increase the complexity of pre- and post-processing. Part of this overhead lies within the domain of hardware development, e.g., qubit reset and readout. Much of this is, however, classical computing routines, such as compiling, transpiling, qubit routing optimisation, error mitigation, noise cancelling, to name a few. All of these will become computationally more demanding with increasing qubit count, and will therefore require increasing amounts of classical computing power. Thus, efficiently operating the quantum machines of the future will require an HPC infrastructure in itself, as well as the classical software to run on it.

For reaching quantum advantage as soon as possible, both in general and especially within climate modelling, it is important to develop quantum algorithms keeping the immense, existing classical supercomputing power in mind. This means for example taking full advantage of the available HPC infrastructure for performing the necessary pre- and post-processing of data to and from the quantum machines. For electronic structure problems, as in the case of modelling atmospheric reactions discussed above, HPC resources are needed for providing an initial guess for the quantum computer; in other words, provide the best approximation to the true electronic structure that classical methods can provide, and refine it further on the quantum computer. This exemplifies the need for a broad, multidisciplinary approach to quantum advantage. We need to combine expertise in quantum algorithms, classical HPC algorithms, computer science, AI/ML, and specific domain expertise, also from the end-user side.

11 Sizing Quantum Machines for Climate Modelling

As discussed above, there are several means to achieve quantum advantage for the climate digital twins. Different problems are suitable for different quantum machines and implementations. Efficient quantum solutions for linear systems of equations and computational fluid dynamics have mostly been proposed for gate-based quantum computers. (Hybrid) quantum algorithms for machine learning and quantum chemistry have been proposed for all three major quantum machines classes, quantum annealers, quantum simulations, and gate-based quantum computers. Thus, for climate digital twins, all three classes can potentially be useful. It is important not to focus efforts too narrowly, say, only on gate-based algorithm development. Various approaches should be explored for the different quantum machines, and also combination of approaches, like digital-analog quantum computing and simulation.

We want to emphasise the difficulty in predicting developments of both quantum hardware and software, quantum algorithms. The expectations on hardware are reasonable, but naturally come with large error bars. Still, the progress can be expected to be rather smooth. A much larger uncertainty still comes from the quantum software side, as new discoveries can truly revolutionise the utility of quantum machines. It is completely possible that a novel "Shor's algorithm" for climate modelling will be invented, or that new algorithmic breakthroughs for relaxing the requirements on, say, coherence times for the qubits will be developed. In the case of software, the progress *can* be smooth, but the possibility for (quantum) leaps in efficiency is ever-present.

11.1 Present Day

We are still at the development phase of what is to become a mature supercomputing infrastructure incorporating quantum machines for climate modelling. The quantum hardware is becoming sufficiently stable for performing real-world testing at small scale. Quantum machines are being incorporated with HPC infrastructure in various manners, from cloud access models, via co-located installations, to truly distributed approaches Johansson et al. [2021]. Standards are under development, but not yet in place, for the various components of the full HPC+QC software stack. Different programming models are still developed in parallel. This is a necessary step in the evolution of a fully mature quantum computing infrastructure: we need to try out several different approaches, even through blunt trial-and-error, in order to learn which methods, protocols, and models work best. We have to resist the temptation of unifying different models at a too early stage, even if this complicates the life for hybrid HPC+QC software developers somewhat.

At the time of writing, none of the three platforms, annealers, simulators, or gate-based "universal" quantum computers provide any computational quantum advantage over purely classical methods. All of them have, however, demonstrated the potential for inching closer to the point of quantum advantage for problems resembling real-world use cases. The D-Wave "Advantage" quantum annealer has over 5600 spin qubits. The largest available Pasqal neutral atom simulator presently has 100 qubits, with 200 planned for 2023 and 1000 for 2024. IBM's "Osprey" gate-based QPU has 433 qubits, with a thousand-qubit QPU announced still for 2023.

In recent years, several demonstrations of experiments on quantum machines, performing tasks that would be very hard for classical computers to simulate, have been reported. Keeping in mind that full simulation of sixty fully entangled qubits is already far beyond the capacity of the worlds largest supercomputers, we *should* be entering the regime of potential quantum advantage right now.

Naturally, qubit count is only one figure of merit, as discussed previously. Coherence times, operation fidelity, and qubit connectivity are at least as important; arguably, already around a fifty qubits, these already become more important than the physical qubit count. Presently, the state-of-the-art quantum computers feature roughly a 99.9% fidelity on their operations. This means that on average, one in a thousand operations fail. In order to be able to efficiently suppress errors, a minimum fidelity of 99.9999% is required Google Quantum AI [2023]. We are thus presently roughly three orders of magnitude below target on fidelity, one of the main reasons why useful quantum advantage has not yet been demonstrated.

11.2 3-5 years

Incremental advantage, or at least a convincing prospect of advantage over purely classical methods relevant to ClimateDT should by this time be exhibited. The qubit count and quality of all three classes of quantum machines should be such that reasonably reliable estimates of what types of algorithms can be expected to show true quantum advantage can be made.

The availability of quantum computers will still be so low, that for production-scale climate modelling, purely classical HPC infrastructure will be used. Benchmarking of hybrid HPC+QC methods for relevant machine learning and electronic structure tasks will be underway, paving the way for models based at least partly on data produced by quantum machines.

11.3 15 years

By this time, all surviving quantum machine classes would have sizes in the hundred-thousand to a million physical qubit regime. This implies a logical qubit count of at least a hundred, possibly thousands, and would be sufficient for executing sufficiently complex simulations or circuits for being directly relevant for climate modelling.

The manner in which the quantum machines will be utilised depends heavily on the progress in "clock speed": how fast can the Hamiltonian evolution be driven, how short are the gate-operation times, and so on. We estimate that execution speed will still be much too slow for incorporating quantum operations directly inside the workflow of the digital twin, due to the strict and short wall-time requirements imposed on each time step, which at this time would be counted in milliseconds.

It is possible that on-chip QPU technology will be implemented within 15 years, although we estimate that sufficiently mature solutions would be some time away still. With quantum and classical processing on the same chip, it may be possible to seamlessly execute short quantum-accelerated subroutines as part of a more complex calculation. This would require efficiently integrated and error-corrected/mitigated quantum coprocessing technology, in an analogous manner to floating-point units (FPUs) of standard CPUs. Tightly integrated QPUs would have a major impact on time-step evolution problems, like ClimateDT, as they would enable quantum acceleration within the individual time steps.

With a million physical qubits, or a thousand logical ones, highly accurate modelling of atmospheric reactions of small molecules exhibiting complex electronic structures will be possible, which will increase the accuracy of the digital twins notably. We note that atmospheric chemistry is a prime candidate for early advantage from quantum-accelerated electronic structure solutions, due to the relatively small size of the molecules involved; the number of electrons that need to be described is very moderate compared to, say, enzymatic reactions. The atmospheric reaction flows can then be included in the digital twins in a parameterised manner. In the same vein, quantum-assisted machine learning will be used for training models relevant for various parts of the digital twins that now rely on machine learning for speed and/or accuracy.

12 SWOT analysis

12.1 Strengths

- The climate modelling community has a deep understanding of the problem at hand, and the bottlenecks present, both from the efficiency and accuracy points-of-view
- A recognised high-priority problem: resources available for finding solutions

12.2 Weaknesses

- Understanding of the applicability of quantum computing to climate modelling limited
- Quantum-acceleration presently not seen as a viable route, due to the "big data" nature of digital twins

12.3 Opportunities

- Progress in QC hardware and software capacity can enable more accurate models
- Global drive for supporting hybrid HPC+QC software development

12.4 Threats

- Development of sufficiently powerful QC hardware/software delayed
- Lack of long-term funding commitment to development, in case near-term gains do not live up to (inflated) expectations.

13 Conclusions

In their present form, digital twins of the climate are largely not amenable to quantum acceleration, due to their reliance on large amounts of both input and output data, and very short wall-time of the individual time steps. Despite this, quantum computers have the potential to both speed up current climate digital twins, as well as increase their accuracy. The increase in accuracy follows both from enabling higher resolution of the digital twins, and from the possibility of setting up a more complete physical model of the Earth and the dynamic processes that govern the time-evolution of the climate. The build-up towards computational quantum advantage is steady. In the medium term, within the next ten years, quantum computing can provide incremental but notable improvements to the accuracy of the models.

In the longer term, with fault-tolerant, or at least near-FT quantum machines, the gains can be of decisive importance. For computational quantum advantage in climate modelling, we estimate that this will require at least 15-20 years of further hardware development. Digital twins of the Earth and climate will thus not be among the first applications where notable quantum advantage will be found, but in time, also climate modelling will experience a quantum revolution.

Alongside the hardware, it is crucial to actively develop a diverse software ecosystem around the maturing quantum machines. The problems and subtasks that make up ClimateDT will need to be reformulated in a suitable manner in order to be amenable to quantum computing. The new quantum software will need to be seamlessly integrated with existing workflows based on classical processing. Supporting software, such as compilers, error mitigation and error-correction routines, even programming languages themselves need to be created.

Even if we would have sufficiently mature and powerful quantum machines right now, they would be practically useless for climate modelling, due to the lack of software. Just preparing the software framework for efficiently running a climate digital twin partly on quantum machines can take a decade. The Destination Earth initiative itself is about making existing classical software run efficiently on the latest supercomputers. In essence, this largely means switching the code base from CPUs to GPUs, from one classical computing platform to another. Switching from classical to quantum is a whole different level of hard. Therefore, we need to start the transition now, in order for the software to be ready when the hardware is. If and when this happens, it will boost information-based climate adaptation efforts significantly.

The software development will most likely require dedicated non-commercially motivated funding. Long-term commitment is needed, and the problem to be solved, creating a highly reliable and efficient long-term climate model, has little direct economic impact; much of the missing work is still fundamentally basic research. The indirect impact on economy and society as a whole is, of course, immense.

It is crucial to recognise the massive computational power of already existing high-performance computing (HPC) infrastructure. Direct replacement of classical methods by analogous quantum methods will not bring significant speedup. We cannot just recompile existing classical subroutines in the ESMs to run on quantum machines and expect any advantage. Instead, quantum computing needs to augment and enhance present modelling procedures. By approaching the actual problem from new angles, presently used classical subtasks and approximations can also be rendered obsolete. Quantum computing is *different*, and therein lies its strength. By solving old problems in a new way, or enabling solutions to previously intractable problems, quantum machines can accelerate HPC infrastructure in a meaningful way, complementing binary supercomputers.

Here, we have identified two main approaches for gaining quantum advantage for the climate DTs: quantum-enhanced machine learning approaches and the extension of the models to include presently missing physics and chemistry, such as atmospheric reactions. These are naturally just two examples of latent quantum advantage. Many more are expected to be uncovered by dedicated efforts. The futures of HPC and climate digital twins are quantum-accelerated; how far ahead that future lies, depends on the combined efforts of several fields of science and technology.