Climate-change modelling at reduced floating-point precision with stochastic rounding

Tom Kimpson | E. Adam Paxton | Matthew Chantry | Tim Palmer

University of Oxford, Oxford, UK

Correspondence
T. Kimpson, University of Oxford, Oxford, UK.
Email: tom.kimpson@physics.ox.ac.uk

Funding information
H2020 European Research Council, Grant/Award Number: 741112

Abstract
Reduced-precision floating-point arithmetic is now deployed routinely in numerical weather forecasting over short timescales. However, the applicability of these reduced-precision techniques to longer-timescale climate simulations—especially those that seek to describe a dynamical, changing climate—remains unclear. We investigate this question by deploying a global atmospheric, coarse-resolution model known as Simplified Parameterizations Primitive Equation Dynamics (SPEEDY) to simulate a changing climate system subject to increased CO₂ concentrations, over a 100-year timescale. Whilst double precision is typically the operational standard for climate modelling, we find that reduced-precision solutions are sufficiently accurate. Rounding the floating-point numbers stochastically, rather than using the more common “round-to-nearest” technique, improves the performance of the reduced-precision solutions notably. Over 100 years, the mean bias error (MBE) in the global mean surface temperature (precipitation) relative to the double-precision solution is $+1.8 \times 10^{-2} \text{ K} (-8 \times 10^{-4} \text{ mm}(6 \text{ hr})^{-1})$ when integrating numerically at half precision (10 significant bits) with stochastic rounding. By examining the resultant climatic distributions that arise after 100 years, the difference in the expected value of the global surface temperature relative to the double-precision solution is $\leq 5 \times 10^{-3} \text{ K}$ and that for precipitation is $8 \times 10^{-4} \text{ mm}(6 \text{ hr})^{-1}$. Whilst further research is necessary to extended these results to more complex and higher-resolution models, they indicate that reduced-precision techniques and stochastic rounding could be suitable for the next generation of climate models and motivate the use of low-precision hardware to this end.

Keywords
climate models, numerical modelling, reduced-precision techniques, SPEEDY
1 | INTRODUCTION

The use of reduced-precision floating-point arithmetic in place of the conventional 64-bit double precision (i.e., IEEE-754 Float64) has gained recent attention in numerical weather modelling (Chantry et al., 2019; Hatfield et al., 2019) for the potential to reduce computational demands on otherwise intensive processes. This saving in computational expense can then be reinvested in, for example, increasing model resolution, hence improving the accuracy of forecasts (Lang et al., 2021). For this reason, reduced-precision methods are now used in operational weather forecasts: for example, the European Centre for Medium-Range Weather Forecasts runs the atmospheric component of the Integrated Forecast System (IFS) in single (32-bit float) precision (Vana et al., 2016; Maass, 2021). Weather forecasting centres in both the UK (Gilham, 2018) and Switzerland (Rudisuhl et al., 2014) have also explored the benefits of reduced-precision arithmetic. Beyond numerical weather forecasting, reduced-precision methods are now deployed routinely in neural network models (Hopkins et al., 2020; Gupta and Ranga, 2021; Rehm et al., 2021; Noune et al., 2022), which motivates the development of dedicated CPU and accelerator hardware and architectures (TensorFlow, 2020; Training With Mixed Precision, 2021) that can take advantage of reduced-precision formats. These advances in hardware may in turn also be exploited by the numerical weather modelling community.

Given the success of reduced-precision methods in numerical weather prediction and the emergence of corresponding hardware, it is also of interest to investigate the use of the same techniques over longer timescales, from weather to climate modelling. This question was first investigated in Paxton et al. (2022) (hereafter Pax21), who explored the rounding error introduced as a result of using reduced precision for a global atmospheric climate model. They demonstrated that the rounding error using 12 significand bits was negligible and did not disrupt the long-term climatology, whilst model accuracy could be maintained with acceptably minor errors emerging at 10 significand bits. They also briefly consider alternative rounding modes, beyond the standard deterministic “round-to-nearest”. In particular, the introduction of stochastic rounding appeared to have a modest effect to partially mitigate the rounding errors introduced at reduced precision.

Motivated by these results, in this work we build on the previous study of Pax21 to investigate the efficacy of reduced precision when applied to modelling a changing climate system, subject to increased CO2 levels. The global atmospheric modelling of Pax21 used a constant-in-time sea-surface temperature (SST) anomaly field, such that the boundary conditions were annually periodic and the system admitted a well-defined, invariant climate distribution. Whilst this enabled them to obtain more stringent bounds on the rounding error through minimizing any external variability due to, for example, nonstationarity, for the purposes of applying reduced-precision arithmetic to accurate climate models it is important to consider a changing climate, particularly in response to variations in atmospheric CO2.

In this work we will deploy the same global atmospheric model as used in Pax21, but now remove these previous constraints and instead use a nonstationary SST anomaly field along with an instantaneous quadrupling of the atmospheric CO2. This enables us effectively to perform a climate-change experiment, where we will focus on two key questions:

1. Can we reproduce the climate-change signal at reduced floating-point precision?
2. Can stochastic rounding help mitigate any errors incurred at reduced float precision?

The atmospheric modelling used here is an inherently simplified version of a full state-of-the-art climate model. However, definitive answers to the above questions using this simplified model would give strong indications as to the potential of reduced-precision arithmetic and alternative float rounding methods when applied to more complex climate simulations, and may suggest fruitful avenues for increasingly accurate and computationally demanding climate modelling.

The article is organised as follows. In Section 2 we give a short overview of round-to-nearest and stochastic rounding techniques for finite-precision arithmetic. In Section 3 we describe the atmospheric model used and our climate-change experiment in more detail, before running the experiment at different float precisions. We explore both the transient behaviour as the climate is changing and the ultimate invariant distribution (i.e., climate) that the solutions equilibrate to. A general discussion on the experiment results and potential implications is presented in Section 4, before concluding remarks and suggestions for future directions of work.

2 | ROUNDING FLOATING-POINT NUMBERS

In general, real numerical values are not exactly representable using a finite-precision arithmetic number system, such as is used on a computer. Instead, it is necessary to map—or round—from the true number to one that can be represented by the computer. This operation...
introduces a rounding error— the difference between this true result and the representable value. The IEEE 754 standard (IEEE, 2019) uses a “round-to-nearest” (RN) mapping where the operation result is rounded to the nearest representable value. It is fundamentally a deterministic process; a given result will always be mapped to the same representation in the same floating-point system. The error introduced by RN can be shown to satisfy (Higham, 2002; Connolly et al., 2021)

\[ R_{\text{RN}}(x) = x(1 + \delta), \quad \delta \leq \varepsilon / 2, \]

where \( \varepsilon \) is the machine epsilon.

Rather than deterministic RN, it is also possible to use an alternative method known as stochastic rounding (SR). In this case, the result of an arithmetic operation is mapped to a representable form as

\[ R_{\text{SR}}(x) = \begin{cases} 
\lfloor x \rfloor & \text{with probability } p(x), \\
\lceil x \rceil & \text{with probability } 1 - p(x),
\end{cases} \]

(\text{Forsythe, 1959; Hull and Swenson, 1966}). As noted in Connolly et al. (2021), there are some standard properties that hold under RN which no longer hold under stochastic rounding. For instance, under SR, generally \( R_{\text{SR}}(-x) \neq -R_{\text{SR}}(x) \) (lack of rounding symmetry due to stochasticity) and \( x \leq y \Rightarrow R_{\text{SR}}(x) \leq R_{\text{SR}}(y) \) (nonmonotonicity of rounding). Moreover, the rounding error now satisfies

\[ R_{\text{SR}}(x) = x(1 + \delta), \quad \delta \leq \varepsilon, \]

a less stringent bound than for RN. Furthermore, since stochastic rounding naturally requires random number generation, it is likely more expensive than a deterministic process.

In spite of these apparent drawbacks, SR has some important properties which make it advantageous from the perspective of both training neural networks (Gupta et al., 2015; Na et al., 2017; Wang et al., 2018; Xia et al., 2021) and finding numerical solutions to differential equations (Hopkins et al., 2020; Fasi and Mikaitis, 2021; Croci and Giles, 2022). In particular, SR is a statistically unbiased rounding method, since \( E(R(x)) = x \). This also helps prevent numerical stagnation, whereby, when updating a relatively large number \( x \) by some small quantity \( \delta \) using round-to-nearest, \( E(R(x + \delta)) = x \). Consequently the information carried in \( \delta \) is lost, whereas it is preserved under SR. To take a relevant physical example, if, during numerical climate modelling, a temperature tendency is described incorrectly as constant (stagnant) when using deterministic rounding, this will affect the melting of permafrost, for example. A further advantage of SR is that the rounding errors are uncorrelated, being mean-independent, zero-mean random variables (Croci and Giles, 2022). Consequently, the global rounding error is smaller than in the RN case. Whilst stochastic rounding has not traditionally been supported in hardware, the emerging trend of low-precision arithmetic in machine learning has strongly motivated the development of new hardwares. For example, the new Graphcore Intelligence Processing Unit (IPU)—developed explicitly for machine learning applications—has inbuilt stochastic rounding support (Jia et al., 2019; IPU Hardware Overview, 2022). A comprehensive review of SR hardware that is available and in development is presented in Croci et al. (2022).

3 | CLIMATE-CHANGE EXPERIMENT

3.1 | Overview

Throughout this work, following Pax21, we will use a global atmospheric model known as Simplified Parameterizations PrimitivE Equation Dynamic, version 41 (SPEEDY: Molteni, 2003) that has been modified to run in low precision (Saffin et al., 2020) utilizing a reduced-precision emulator RPE v5 (Dawson and Düben, 2017). SPEEDY is an intermediate-complexity, coarse-resolution model that uses a spectral transform dynamical core with a spectral resolution of T30, a 96 × 48 Gaussian grid, a 40-minute timestep, and eight vertical sigma levels. The numerical integration timestep itself uses a leapfrog scheme in conjunction with a Robert–Asselin–Williams filter (Williams, 2009). Daily updated boundary conditions are determined by taking monthly averaged values (e.g., for sea ice) and then linearly interpolating. We use the same set of initial conditions as described in Pax21.

We are fundamentally interested in being able to model a changing climate. A widely used metric in climate modelling is the equilibrium climate sensitivity (ECS), which describes the change in the global mean surface temperature after an instantaneous doubling of CO2 compared with pre-industrial levels. In order to evaluate the sensitivity of a climate model to CO2 forcings and determine the ECS, it is common practice to use a “4xCO2” experiment, where a spun-up pre-industrial run is subject to an abrupt quadrupling of the atmospheric CO2 concentration; the initial Diagnostic, Evaluation and Characterization of Klima (DECK) requirements for participation in Phase 6 of the Coupled Model Intercomparison Project...
(CMIP) include a 4xCO2 experiment (Eyring et al., 2016), whilst a previous EC-Earth experiment has explored the change in the ECS when progressing from CMIP5 to CMIP6 models (Wyser et al., 2020). For our climate-change experiment, we follow the same example and instantaneously quadruple the CO2 concentration. Specifically, we set the absorptivity of air in the CO2 band to be 21, as a rough estimate of a value four times the pre-industrial level. This quadrupling is advantageous for our purposes, since it subjects the reduced-precision arithmetic to a more strenuous test than simply doubling the CO2 concentration. Rather than holding the SST field constant as in Pax21, we take the SST from the corresponding EC-Earth experiment, regridded to match the SPEEDY grid. As a shorthand we will refer to this system with nonconstant SST and increased CO2 concentrations as a “4xCO2” world.

We run the experiment at the following precisions, where we truncate the significand bits only:

- “Double”—53 significand bits, 11 exponent bits,
- “Single”—23 significand bits, 11 exponent bits,
- “Half”—10 significand bits, 11 exponent bits.

For the solutions at half precision we will consider both round-to-nearest and stochastic rounding methods. We will use the notation X.Y to refer to a solution with X significand bits using a rounding method Y: for example, 10_SR labels the half-precision solution, using stochastic rounding. We note that our definitions of “Single” and “Half” are distinct from the IEEE-754 definition, since we do not modify the number of exponent bits, being concerned primarily with numerical precision rather than dynamic range. Throughout this work, we will take “Single” and “Half” to be defined as above, and reserve “Float32” and “Float16” to refer to the IEEE-754 definitions.

3.2 Transient effects of climate change

Initially, consider as a demonstrative example a particular SPEEDY integration over 100 years from a single initial condition, for the 4xCO2 world. The evolution of the global mean surface temperature (i.e., the latitude-weighted global average of the surface temperature at every grid point) over a 100-year timescale is presented in Figure 1, along with the error relative to the full double-precision solution, which we take as our ground truth. We can see that, for a 4xCO2 world, the global mean surface temperature generally increases logarithmically by just under ~4 K over 100 years. Against the background
of this general increase, there is also a shorter timescale, $O(\text{years})$, oscillation in the temperature. Both the overall profile and the magnitude of the true (Float64) solution are generally reproduced well by all the reduced-precision solutions. Whilst there is clearly a high degree of volatility year-on-year in the error, the relative accuracy in the reduced-precision solutions can be discerned clearly. Single precision exhibits the best performance, tracking the true solution accurately and without bias: averaging over the 100-year period, the mean bias error (MBE) is $-3.6 \times 10^{-3}$ K. The half-precision solution using round to nearest is less well-performing, with an MBE approximately 12 times that of single precision. This is evident from Figure 1, where the 10_RN solution is distinctly offset from the true time series, consistently underestimating the surface temperature whilst still following the general profile. However, the addition of stochastic rounding to the half-precision solution improves the performance notably, reducing the bias relative to Float6 by around a factor of 4. Similar results are observed if we instead consider global mean precipitation (Figure 2); the double-precision solution is reproduced well at both single precision and half precision with stochastic rounding, where the MBE for both is highly comparable ($\sim -8 \times 10^{-4}$ versus $\sim -7 \times 10^{-4}$ mm·(6 hr)$^{-1}$). In contrast, the half-precision solution using round to nearest is distinctly less accurate, by orders of magnitude (Table 1).

From these single runs, a clear hierarchy of solutions emerges, with 23_RN > 10_SR > 10_RN. Whilst the 10_SR solution is not quite as good as the single-precision solution, it still exhibits a generally improved performance over the 10_RN solution, which has an identical number of information bits. Moreover, the 10_SR solution is broadly comparable with the 23_RN solution, despite carrying 13 fewer bits of significand information. Furthermore, the computer time saved by operating at a reduced precision could be reinvested to, for example, increase the vertical resolution, which could have a stronger effect on improving the accuracy than the small degradation.

**TABLE 1** Mean bias error (MBE) in temperature and precipitation, averaged over 100 years, for each of the alternative number formats, cf. Figures 1 and 2.

| Number format | $\text{MBE}_{\text{temperature}}$ (K) | $\text{MBE}_{\text{precipitation}}$ (mm·(6 hr)$^{-1}$) |
|---------------|-------------------------------------|---------------------------------|
| Single        | $-3.6 \times 10^{-3}$               | $-8 \times 10^{-4}$             |
| Half          | $-4.4 \times 10^{-2}$               | $-1.2 \times 10^{-2}$           |
| Half + SR     | $1.4 \times 10^{-2}$                | $-8 \times 10^{-4}$             |
of 10_SR. This is an encouraging initial result with regard to the two questions we posed at the start of this article; the climate-change signal can be reproduced accurately at low precision and stochastic rounding generally improves the solution—at least for this demonstrative single run using a simplified atmospheric model.

Building upon this single run, we can now consider an ensemble of initial conditions and explore the transient behaviour across all ensemble members. We use five members per ensemble for each precision. The results are presented in Figures 3 and 4 for temperature and precipitation, respectively. The behaviour is generally similar to our previous integration from a single initial condition; all the reduced-precision solutions generally perform well, reproducing the true double-precision solution faithfully to within $\sim 0.1$ K for temperature and $\sim 0.015$ mm-$\text{hr}^{-1}$ for precipitation. The same hierarchy of solutions is also evident, with the half-precision stochastic rounding solution outperforming round-to-nearest; for temperature, stochastic rounding results in a solution that is approximately twice as accurate as the round-to-nearest, whilst for precipitation it is $\sim 14$ times more accurate. It is also notable that in both cases the 10_RN profile generally underestimates the true solution, which is characteristic of numerical stagnation. For temperature, this underestimation is especially prominent at early times, when the global surface temperature is increasing rapidly due to the changing climate. The generally logarithmic nature of the true temperature profile as the rate of warming decreases as the climate starts to equilibrate dampens the impact of this stagnation, giving the 10_RN solution a chance to “catch up” at later times. This then suggests that using the 100-year average of the error as an accuracy metric is perhaps too generous to the 10_RN solution, which does not follow the true solution accurately over all time in the way that 23_RN and 10_SR do, but instead has its score buoyed at later times. For instance if we calculate the MBE over the initial 25 years rather than the full 100-year integration, then the MBE for 10_RN is $7.2 \times 10^{-2}$ K, while for 10_SR it is $4.9 \times 10^{-3}$ K. Conversely, the precipitation is not subject to such rapid increases and the accuracy of the 10_RN solution decays gradually over time. Whilst the 10_SR solution performs well and can be seen to be superior to 10_RN, it is worth acknowledging that it is generally slightly less performant than 23_RN for both temperature and precipitation. The 10_SR and 23_RN solutions are highly comparable for precipitation, with 23_RN a touch more accurate, whilst the superiority of 23_RN over 10_SR is much more evident for temperature. However, 10_SR achieves almost on-par performance over 100 years with

**FIG URE 3** Time evolution of the global mean surface temperature over 100 years for a 4×CO2 world, averaged over five ensemble members (top panel), and the ensemble bias error (bottom panel). The ensemble temperature dispersion is present, but too small to be seen on this scale. The ensemble bias is presented as a decadal average so as to smooth out the volatility and display the general trends. [Colour figure can be viewed at wileyonlinelibrary.com]
FIGURE 4  As Figure 3, for global mean total precipitation. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2  As Table 1 for an ensemble of runs, cf. Figures 3 and 4.

| Number format | MBE_{temperature} (K) | MBE_{precipitation} (mm (6hr)}^{-1} |
|---------------|-----------------------|--------------------------------------|
| Single        | $2 \times 10^{-4}$    | $8 \times 10^{-5}$                    |
| Half          | $-3.5 \times 10^{-2}$ | $-1.1 \times 10^{-2}$                |
| Half + SR     | $1.8 \times 10^{-2}$  | $-8.2 \times 10^{-4}$                |

13 fewer bits of significand information, and indeed 43 fewer significand bits than the double-precision solution! More broadly, the strong performance of single precision suggests that, similar to numerical weather modelling, climate models could potentially be run straightforwardly at Float32 (Table 2).

In addition to considering the global average time series, we can also investigate the spatial behaviour of the reduced-precision solutions explicitly: are there particular geographic areas where 10_RN is failing but 10_SR succeeds? In order to do this, we compare our “competitor” ensembles at each precision against a separate control ensemble, which is composed of five members at double precision. Figures 5 and 6 show the spatial distribution of the difference between the competitor and control ensembles, averaged over 100 years, for temperature and precipitation, respectively. In both cases, the 10_SR outperforms 10_RN globally. For the surface temperature, the 52_RN solution is relatively uniform on this ~K scale, showing that the inherent chaotic variability in the mean surface temperatures is sub-Kelvin. The 10_SR map is also relatively featureless, and is highly comparable with the double-precision solution. Conversely, the 10_RN solution shows large overestimates relative to the control group, mainly over land, primarily in North Africa and the Middle east as well as North America and South Africa, and large underestimates over the sea. For precipitation, the same trend regarding the relative performance of 10_SR and 10_RN is true, but now the errors are localised to the regions with the largest rates of precipitation, namely the Tropics. In this case, neither solution exhibits a particularly strong over/underestimate bias in particular regions.

3.3  Invariant climate distribution

The preceding analysis was concerned with the transient effects of climate change. During this time, the climates are ill-defined and it is not possible to define consistently some invariant distribution that can then be compared at different float precisions—the approach used in Pax21 for constant SSTs and no external forcings. However, it
is of interest to be able to examine the invariant distribution that arises over time, that is, the climate settled upon after climate change has finished. To this end, we now take the system after 100 years of climate change and re-impose constant SST fields, analogous to Pax21, and integrate for a further 10 years. We will construct a control ensemble \( e^{\text{control}} \) at double precision and a competitor ensemble \( e^{\text{competitor}} \) at each of the double/single/half float precisions. We use the same set of initial conditions \( i_j \) for \( j = 1, \ldots, 10 \) as described in Pax21. These initial conditions were obtained via integrating a system from rest for 11 years at double precision with annually periodic boundary conditions, discarding the first year, and taking the start of each subsequent year as the \( j \)th initial condition. The control ensemble is constructed from the set of initial conditions \( i_1 \ldots i_5 \), whilst a competitor ensemble is constructed from \( i_6 \ldots i_{10} \). We will use the notation \( e_j \) to refer to the \( j \)th member of an ensemble, which has initial condition \( i_j \).

In order to compare the control and competitor ensemble probability distributions we adopt as our key metric the Wasserstein distance (WD). The WD defines a distance between two distributions intuitively as the optimal cost, with respect to some cost function \( c(x,y) \), of transporting a probability mass from position \( x \) to position \( y \). The
FIGURE 7 Area-weighted mean grid point Wasserstein distances for (a) temperature and (b) precipitation for the competitor ensembles $e_{\text{competitor}}$ relative to the control ensemble $e_{\text{control}}$, for constant SST fields over 10 years, after 100 years of changing climate integration. The WD can be considered as an upper limit on the difference in the expected value between two distributions from Equation (8). [Colour figure can be viewed at wileyonlinelibrary.com]

$p$-Wasserstein distance between two distributions $\mu$, $\nu$ is

$$W_p(\mu, \nu) = \left( \inf_{\gamma \in \Gamma(\mu, \nu)} \int c(x,y)^p \, d\gamma(x,y) \right)^{1/p},$$

(5)

where $\gamma(x,y)$ is the transport plan, $\Gamma(\mu, \nu)$ is the set of all couplings of $\mu$ and $\nu$, and $c(x,y)$ is the cost function. It is well-known (Dudley, 1969) that the Wasserstein distance suffers from the curse of dimensionality for dimensions $d \geq 3$:

$$\mathbb{E}[W_1(\mu_n, \mu)] \approx n^{-1/d}.$$  

(6)

To sidestep this issue, we marginalize on to the distributions spanned by the individual gridpoints and take the WD between these one-dimensional distributions. Such an approach is also adopted in Paxton et al. (2022) and Vissio et al. (2020). We also take $p = 1$ and $c(x,y) = |x - y|$, such that the WD values quoted in this work are given by

$$W_1(\mu, \nu) = \inf_{\gamma \in \Gamma(\mu, \nu)} \int |x - y| \, d\gamma(x,y).$$

(7)

From the perspective of climate modelling, the WD has a nice physical interpretation given by the Monge–Kantorovich duality:

$$|\mathbb{E}(X_\mu) - \mathbb{E}(Y_\nu)| \leq W_1(\mu, \nu),$$

(8)

that is, the difference in the expected value of some random variable between two distributions is no bigger than the WD value itself. More explicitly, if the WD for surface temperature between two distributions is 1 K, then this tells us that the difference in the expected temperature between the distributions is no greater than 1 K. For a full review of the Wasserstein distance, including favourable properties and comparison with other metrics, we refer the reader to the Appendix of Pax21.

The mean Wasserstein distances (area-weighted average over all grid points) can be seen in Figure 7 for both the global surface temperature and the total precipitation. For both variables, the distribution results quantified by the Wasserstein distance mirror the transient time-series effects that we demonstrated in the previous section (e.g., Figure 1) Firstly, the single-precision solution reproduces the double-precision solution with high fidelity. The Wasserstein distance error (i.e., the difference between the double solution and the reduced-precision solution) after 10 years is $4 \times 10^{-3}$ K for temperature and $5 \times 10^{-4}$ mm $(6 \text{hr})^{-1}$ for precipitation. Secondly, the half-precision solution with round-to-nearest, despite being our worse performing solution, exhibits generally reasonable performance, approximating the double-precision solution well with WD errors of 0.1 K and 0.012 mm $(6 \text{hr})^{-1}$ for temperature and precipitation, respectively. Again, the addition of stochastic rounding to the half-precision solution improves the performance vastly; for temperature, the WD error is $5 \times 10^{-3}$ K, and for precipitation, $8 \times 10^{-4}$ mm $(6 \text{hr})^{-1}$, highly comparable with the single-precision behaviour. To emphasize the interpretation of the Wasserstein distance, which may
be unfamiliar to readers, this means that, after 110 years of numerical integration (100 with time-variable SSTs, 10 with time-constant), the difference in (for example) the expected value of the climate global mean surface temperature between the double-precision solution and the half-precision solution with stochastic rounding is \( \leq 5 \times 10^{-3} \) K. Referring back to the two original questions that we posed at the start of this work, we can see that the climate-change signal can be reproduced well at reduced precision (highly accurate at single precision, small errors at half precision) and that the errors we see at half precision can be reduced significantly by the addition of stochastic rounding (half precision with stochastic rounding is comparable with double precision).

4 | DISCUSSION

Since, fundamentally, stochastic rounding trades short-term accuracy for long-term statistics, it seems reasonable to suggest that it may lend itself favourably to long-timescale climate simulations at reduced precision. This is indeed what we have observed; whilst a naive reduction from double precision to half precision using deterministic rounding results in a solution that is generally accurate with some small errors, the addition of stochastic rounding improves the performance significantly, to be almost on parity with the full Float64 solution. This definitively answers the two questions we posed in the Introduction at the start of this work. Whilst it is difficult to extrapolate from the simple atmospheric model we have used to more complex state-of-the-art and next generational models, the strong performance at reduced precision suggests that the use of reduced precision in more complex simulations could be a fruitful avenue to explore, at least in some parts of the model. At higher resolutions, it could be the case that the derivatives are too small to be represented naively at reduced precision. In this case, it would be necessary to either scale the values into an appropriate range or else perform the computations at high precision (e.g., Float32), and then round the numbers stochastically to 16 bits before they are transported to a different part of the computer. Indeed, most computing energy in a modern supercomputer is spent transporting data from one node to another or to memory (Kogge and Shalf, 2013). By using stochastic rounding for transport rather than calculating derivatives explicitly, we might be able to get the advantages of reduced precision in terms of savings of machine energy, but without having to degrade the calculation of derivatives. SR is also particularly advantageous for application on these large computers, since any computational cost incurred is likely to be dwarfed by the expense of transporting data and so the cost of 10_RN and 10_SR would be highly comparable. Whilst this study has focused exclusively on issues of numerical precision by modifying the number of significant bits, a climate model implemented at, for example, Float16 would also have to deal with potential issues relating to dynamic range, requiring either mixed precision or scaling. Some hybrid number formats, for example, TensorFloat-32 (Paresh, 2022), are now being developed, which combine the precision of Float16 with the dynamic range of Float32 for specific tensor operations.

From our experiments, it is impossible to predict the exact computational savings that could be made for large-scale climate models from using reduced-precision number formats, since we are working with an emulator that runs in software, rather than directly on hardware, and so incur additional overhead that makes these calculations more costly. Any quantitative estimates on the cost reduction must also be caveated by considerations of, for example, model resolution, implementation, and choice of computer architecture or compiler. These estimates are complicated further if different parts of the model use different number formats. However, there have been a number of previous works that have tried to make savings estimates. For instance, simulations of Lorenz ’96 with reduced precision (32 bits) suggest power savings of 49% in the adder–subtractor circuit and 75% in the multiplier circuit (Düben et al., 2014). Additionally, running the forecasting model of the IFS at 64 rather than 32 bits leads to a 40% reduction in the runtime (Vana et al., 2016). Similarly, the Nonhydrostatic Icosahedral Grid Atmospheric Model (NICAM) run at single precision exhibits a 46% reduction in the runtime (Nakano et al., 2018), whilst the Surface Quasi-Geostrophic Model, an idealised atmospheric model, at single precision displays potential savings of 76% (Thornes et al., 2017). More theoretically, for simulations such as climate simulations, which are heavily limited by data transport or memory, then the cost savings should be in direct proportion to the number of bits carried.

In general, it is difficult to identify all precision bottlenecks in a model, especially for those models that are larger and more complex. An example bottleneck that we pinpointed in the SPEEDY model where the low-precision RN solution starts to incur larger errors which are mitigated by using stochastic rounding is in the interpolation of the monthly climatological sea–land fields down to a daily cadence. For simpler and smaller codes like SPEEDY, it is possible to identify such precision bottlenecks. In this case, one could rescale the relevant equations to ensure that the float operations can be described by the necessary number format as regards dynamic range and precision (see, e.g., Klöwer et al., 2022). Conversely, for larger, more complex codes, rescaling equations and
refactoring code can be prohibitively difficult. Instead, as evidenced in this work, it could be more straightforward to use stochastic rounding as a drop-in replacement for deterministic rounding at reduced float precision. Indeed, the Graphcore Colossus Mk2 IPU chips (IPU Hardware Overview, 2022), which implement stochastic rounding at the hardware level, allow for just such an approach, where the user can simply swap out deterministic for stochastic rounding. If other chip makers follow the same example, it will become much more feasible to implement reduced-precision stochastic rounding on large-scale models. More generally, inefficiencies or small bugs that might be overlooked or gotten away with when using double precision can become important at reduced precision. An apparent failure of the model when going to reduced precision is then not necessarily due to some fundamental loss of information, but can instead be due to the underlying code being structured in a way that is suboptimal for reduced-precision methods. Whilst this does introduce an additional “offline” overhead—that is, one cannot simply change the precision, but may need to optimize the code for that precision—this could be heavily offset by the computational gains that could be made through such optimizations. The development of more modern implementations of global atmospheric models such as SPEEDY will also facilitate the exploration and prototyping of 16-bit climate models. For example, SpeedyWeather.jl (Speedy-Weather.jl, 2022), which is currently under development, is a climate model analogous to SPEEDY, but written in a way that is fully type-flexible to enable arbitrary number formats for performance and analysis simultaneously, with support for hardware-accelerated low precision. This means the model development is precision-agnostic, which allows the common problems of dynamic range and critical precision loss—often incurred from using low-precision number formats—to be addressed directly. Such a model will also be fully differentiable via automatic differentiation, with entire parts of the model able to be replaced with artificial neural networks. This hybrid structure enables an improved representation of climactic processes (by, e.g., training against higher-resolution simulations) and the fitting of models to observational data. The ability to deploy such a model directly on hardware with inbuilt stochastic rounding support would be highly informative with respect to the construction of large-scale 16-bit climate models.

5 | CONCLUSIONS

We have demonstrated that we can model a changing climate system using reduced-precision floating-point numbers, subject to minor errors with respect to the double-precision solution. Whilst the global atmospheric model used in this work is fundamentally an approximation to higher-order state-of-the-art climate models, the effectiveness of reduced-precision techniques motivates the development and further exploration of reduced precision in more complex simulations, at least in some parts of the model. The addition of stochastic rounding improves the half-precision performance significantly, so as to be comparable with the single-precision solution, with only very minor differences from the full double-precision solution. This suggests that the information content of the higher-order bits is low and that significant performance gains could be realised through selective use of reduced precision and stochastic rounding.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant No. 741112).

DATA AVAILABILITY STATEMENT

This implementation of SPEEDY is based off original work by Saffin and Hatfield (Saffin et al., 2020), who developed a Fortran model based on work by Kucharski et al. (2013), modified to include a reduced-precision emulator (Dawson and Dübén, 2017). This model was in turn amended by Pax21 to include stochastic rounding in the emulator (https://github.com/eapax/speedy). The exact SPEEDY model used in this work is based off a branch from Pax21, to which some small refactoring modifications have been made.

ORCID

Tom Kimpson © https://orcid.org/0000-0002-6542-6032
Tim Palmer © https://orcid.org/0000-0002-7121-2196

REFERENCES

Chantry, M., Thornes, T., Palmer, T. and Düben, P. (2019) Scale-selective precision for weather and climate forecasting. *Monthly Weather Review*, 147(2), 645–655.

Connolly, M.P., Higham, N.J. and Mary, T. (2021) Stochastic rounding and its probabilistic backward error analysis. *SIAM Journal on Scientific Computing*, 43(1), A566–A585. https://doi.org/10.1137/20M1334796.

Croci, M., Fasi, M., Higham, N., Mary, T. and Mikaitis, M. (2022) Stochastic rounding: implementation, error analysis and applications. *Royal Society Open Science*, 03, 9.

Croci, M. and Giles, M.B. (2022) Effects of round-to-nearest and stochastic rounding in the numerical solution of the heat equation in low precision. *IMA Journal of Numerical Analysis*, drac012. https://doi.org/10.1093/imanum/drac012.

Dawson, A. and Dübén, P.D. (2017 Jun) rpe v5: an emulator for reduced floating-point precision in large numerical simulations. *Geoscientific Model Development*, 10(6), 2221–2230.
Düben, P., Joven, J., Lingamneni, A., McNamara, H., Micheli, G., Palem, K., and Palmer, T.N. (2014) On the use of inexact, pruned hardware in atmospheric modelling. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 05, 372.

Dudley, R.M. (1969) The speed of mean glivenko-cantelli convergence. The Annals of Mathematical Statistics, 40(1), 40–50 http://www.jstor.org/stable/2239196.

Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J. and Taylor, K.E. (2016) Overview of the coupled model inter-comparison project phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development, 9(5), 1937–1958 https://gmd.copernicus.org/articles/9/1937/2016/.

Fasi, M. and Mikaitis, M. (2021) Algorithms for stochastically rounded elementary arithmetic operations in IEEE 754 floating-point arithmetic. IEEE Transactions on Emerging Topics in Computing, 9(3), 1451–1466.

Forsythe, G.E. (1959) Reprint of a note on rounding-off errors. SIAM Review, 1(1), 66–67. https://doi.org/10.1137/1001011.

Gilham, R. (2018) 32-bit Physics in the Unified Model. United Kingdom: MetOffice. https://digital.nmla.metoffice.gov.uk/download/file/IO_951e52e5-6698-485e-ad33-54d0a2b0ce99.

Gupta, R.R. and Ranga, V. (2021) Comparative study of different reduced precision techniques in deep neural network. Int. J. Mod. Phys. C., 32(2), 2150045.http://dx.doi.org/10.1142/S0129183121500455.

Higham, N.J. (2002) Accuracy and Stability of Numerical Algorithms. 2nd edition. USA: Society for Industrial and Applied Mathematics.

Hull, T.E. and Swenson, J.R. (1966 feb) Tests of probabilistic models for propagation of roundoff errors. Communications of the ACM, 9(2), 108–113. https://doi.org/10.1145/365170.365212.

IEEE Standard for Floating-Point Arithmetic. (2019) IEEE Std 754-2019 (Revision of IEEE 754-2008), 1–84.

IPU Hardware Overview (2022). Accessed April 2022. https://docs.graphcore.ai/projects/ipu-programmers-guide/en/latest/about_ipu.html.

Jia, Z., Tillman, B., Maggioni, M. and Scarpazza, D.P. Dissecting the Graphcore IPU Architecture via Microbenchmarking. arXiv e-prints, arXiv:1912.03413, 2019.

Klöwer, M., Hatfield, S., Croci, M., Düben, P. and Palmer, T. (2022) Fluid simulations accelerated with 16 bits: approaching 4x speedup on A64FX by squeezing ShallowWaters.jl into Float16. Journal of Advances in Modelling Earth Systems, 14(2). https://agupubs.onlinelibrary.wiley.com/action/showCitFormats?doi=10.1029%2F2021MS002684&mobileUi=0

Kogge, P. and Shalf, J. (2013) Exascale computing trends: adjusting to the “New Normal” for computer architecture. Computing in Science and Engineering, 15(6), 16–26.

Kucharski, F., Molteni, F., King, M.P., Farneti, R., Kang, I.S. and Feudale, L. (2013) On the need of intermediate complexity general circulation models: a “SPEEDY” example. Bulletin of the American Meteorological Society, 94(1), 25–30. https://journals.ametsoc.org/view/journals/bams/94/1/bams-d-11-00238.1.xml.

Lang, S.T.K., Dawson, A., Diamantakis, M., Dueben, P., Hatfield, S., Leutbecher, M., Palmer, T., Prates, F., Roberts, C.D., Sandu, I. and Wedi, N. (2021) More accuracy with less precision. Quarterly Journal of the Royal Meteorological Society, 147(741), 4358–4370. https://rmons.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4181.

Maass, C., 2021, Accessed November, 2021. https://confluence.ecmwf.int/display/FCST/Implementation+of+IFS+Cycle+47r2.

Molteni, F. (2003) Atmospheric simulations using a GCM with simplified physical parametrizations. I: Model climatology and variability in multi-decadal experiments. Climate Dynamics, 01(20), 175–191.

Na, T., Ko, J.H., Kung, J. and Mukhopadhyay, S. (2017) On-chip training of recurrent neural networks with limited numerical precision. In: 2017 International Joint Conference on Neural Networks (IJCNN). USA: IEEE, pp. 3716–3723.

Nakano, M., Yashiro, H., Kodama, C. and Tomita, H. (2018) Single precision in the dynamical core of a nonhydrostatic global atmospheric model: evaluation using a baroclinic wave test case. Monthly Weather Review, 146(2), 409–416. https://journals.ametsoc.org/view/journals/mwr/146/2/mwr-d-17-0257.1.xml.

Noue, B., Jones, P., Justus, D., Masters, D. and Luschi, C. 8-bit numerical formats for deep neural networks. arXiv e-prints, arXiv:2206.02915, 2022.

Paresh K. Accessed December, 2022. https://blogs.nvidia.com/blog/2020/05/14/tensorfloat-32-precision-format/.

Paxton, E.A., Chantry, M., Klöwer, M., Saffin, L. and Palmer, T. (2022) Climate modeling in low precision: effects of both deterministic and stochastic rounding. Journal of Climate, 35(4), 1215–1229. https://journals.ametsoc.org/view/journals/jcli/35/4/JCLI-D-21-0343.1.xml.

Rehm, F., Vallecorsa, S., Santeore, V., Pabst, H., Chaibi, A., Codreanu, V., Borras, K., and Krücker, D. Reduced precision strategies for deep learning: a high energy physics generative adversarial network use case. arXiv e-prints, arXiv:2103.10142, 2021.

Rudisuhl, S. and Walser, A. Fuhrer O(2014) COSMO in single precision. Switzerland: Climatology MeteoSwiss.

Saffin, L., Hatfield, S., Düben, P. and Palmer, T. (2020) Reduced Precision Parameterization: Lessons From an Intermediate Complexity Atmospheric Model. Wiley https://eprints.whiterose.ac.uk/158706/.

SpeedyWeather.jl. (2022) GitHub. https://github.com/milankl/SpeedyWeather.jl.

TensorFlow Accelerating AI Performance on 3rd Gen Intel Xeon Scalable Processors with TensorFlow and BFloat16, 2020. Accessed November, 2021. https://blog.tensorflow.org/2020/06/accelerating-ai-performance-on-3rd-gen-processors-with-tensorflow-bfloat16.html.
Thornes, T., Düben, P. and Palmer, T. (2017) A power law for reduced precision at small spatial scales: experiments with an SQG model. *Quarterly Journal of the Royal Meteorological Society*, 10, 144.

Training With Mixed Precision; 2021. Accessed November, 2021. https://docs.nvidia.com/deeplearning/performance/mixed-precision-training/index.html.

Vana, F., Düben, P., Lang, S., Palmer, T., Leutbecher, M., Salmond, D. and Carver, G. (2016) Single precision in weather forecasting models: an evaluation with the IFS. *Monthly Weather Review*, 12, 145.

Vissio, G., Lembo, V., Lucarini, V. and Ghil, M. Ranking IPCC Models Using the Wasserstein Distance. arXiv e-prints, arXiv:2006.09304, 2020.

Wang, N., Choi, J., Brand, D., Chen, C.Y. and Gopalakrishnan, K. Training Deep Neural Networks with 8-bit Floating Point Numbers. arXiv e-prints, arXiv:1812.08011, 2018.

Williams, P. (2009) A Proposed Modification to the Robert-Asselin Time Filter. *Monthly Weather Review*, 08, 137.

Wyser, K., van Noije, T., Yang, S., von Hardenberg, J., O’Donnell, D. and Döscher, R. (2020) On the increased climate sensitivity in the EC-Earth model from CMIP5 to CMIP6. *Geoscientific Model Development*, 13(8), 3465–3474.

Xia, L., Anthonissen, M., Hochstenbach, M. and Koren, B. A Simple and Efficient Stochastic Rounding Method for Training Neural Networks in Low Precision. arXiv e-prints, arXiv:2103.13445, 2021.

**How to cite this article:** Kimpson, T., Paxton, E.A., Chantry, M. & Palmer, T. (2023) Climate-change modelling at reduced floating-point precision with stochastic rounding. *Quarterly Journal of the Royal Meteorological Society*, 149(752), 843–855. Available from: https://doi.org/10.1002/qj.4435