The Thirty-Third International
FLAIRS Conference (FLAIRS-33)

Limits of the Technological Singularity

Curry I. Guinn
Department of Computer Science
University of North Carolina Wilmington
Wilmington, North Carolina, USA
guinnc@uncw.edu

Abstract

The technological singularity hypothesis asserts that the invention of a synthetic intelligence with greater cognitive capacities than a human being will trigger an exponential increase in synthetic cognition and knowledge. Each generation of synthetic intelligences will be able to create new generations of cognitive beings with even greater capabilities than themselves. Some projections envision a future with superintelligences with millions or billions times the cognitive capability of human beings. This paper will argue that the primary function of cognition is to predict the future and make plans based on those predictions. Exponential increases in cognitive capability and knowledge do not necessarily result in exponential increases in the ability to predict and plan for the future. We will show that exponential increases in knowledge may only result in modest linear increases in the ability to predict the future or make plans. Therefore, the actual capabilities of superintelligent machines may only be minimally, if at all, greater than current human capabilities.

The Technological Singularity

The concept of the technological singularity offers a vision of the future where synthetic beings have immense cognitive abilities that are orders of magnitude greater than possessed by human beings. The term "technological singularity" has been popularized by I.J. Good, Vernor Vinge, and Ray Kurzweil (Good, 1965; Vinge 1993, Kurzweil, 1999). The genesis of these theories is that human beings will create a machine with artificial general intelligence with greater cognitive power, problem-solving ability, and creative skills than a human being. In turn, this machine will be able to create synthetic intelligences even greater than its own (since it is more intelligent than its creators). This even more capable intelligence will then be able to design a machine with even greater capacity, and so on. This process will be repeated resulting in superintelligences with exponentially more intelligence than human beings. Some projections envision a single machine having more computational power than all human beings combined by 2050, and the curve does not stop there (Figure 1).

Many prominent futurologists believe this future is plausible and potentially threatening to humanity's well-being (Nick Bostrom, 2014). "'The fact is that AI can go further than humans, it could be billions of times smarter than humans at this point,' Pearson said. (CNBC, 2018)." How could human beings compete with such agents? Would humanity become subservient to these superintelligent machines?

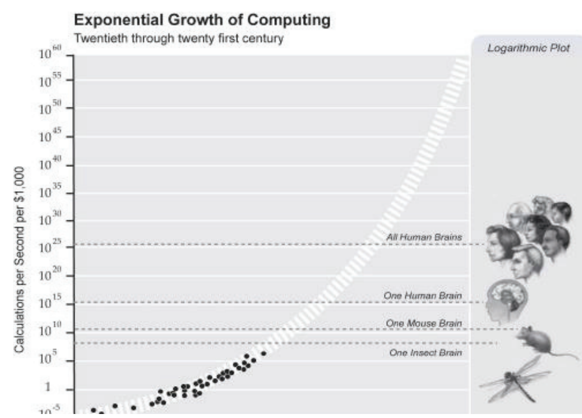

Figure 1 Exponential Growth of Computational Power (Kurzweil, 1999)

In this paper, we will present an approach to cognitive ability that emphasizes its purpose in allowing agents to predict, plan and act in ways that enhance their goals and desires in the future. We will then show that even in a deterministic world, an exponential increase in knowledge may only result in a linear increase in the ability to see into the future. Future events are not completely random; they are possible to predict within certain ranges. However, a substantial increase in knowledge does not necessarily result in narrowing or refining those ranges of possibilities. Therefore, we will conclude that while hypothetical superintelligences may possess substantially more cognitive ability than human beings, they may not be significantly better than human beings at predicting, planning, and acting in order to better control the future.

The Purpose of Cognition

A significant difficulty in discussing the Technological Singularity and superintelligent machines is determining the meaning of intelligence. What does it mean for a being to be more intelligent than another? One approach to this issue is to look at the purpose of cognitive abilities. Cognition in animals provides assistance in the processing and interpretation of perception that allow for the better survivability of the animal. This approach puts a heavy emphasis on the ability to use perceptional and conceptual data to predict the future. "Indeed, the main purpose of cognition seems to consist in anticipatory (offline) preparation: selecting a goal, configuring the system for a particular task, priming goal-relevant action systems, and preparing for the processing of possible trigger stimuli" (Hummel 2015). To describe an agent as having more intelligence than another agent is to say that it is better able to predict the future, anticipate how its possible actions may alter the future, and select the actions that are more likely to lead to reaching the agent's goals. With that approach, let us now address how much better a superintelligence would be at predicting the future as compared to a human being.

Superintelligence and Predicting the Future

The argument for the inevitability of the Technological Singularity is strongly tied to the notion that computational capacity, both in terms of memory and processor speed, will continue to increase exponentially. Much of that argument is based on Moore's Law, the observation that computer capacity/speed doubles every year and a half to two years. Moore's Law has held since the birth of digital computers until today. While there are signs that Moore's Law is beginning to fail, this paper will assume that Moore's Law continues to hold (although the rate of doubling may slow down). Further, the prospect of viable quantum computers suggests that any current slowdown in Moore's Law will be leap-frogged with the advent of commercial quantum computers.

This paper argues that even with that exponential increase in computational ability, these improvements will not result in an exponential increase in the ability to predict and plan for the future. Our argument will involve looking at two limitations on computational ability: 1) the effect of Mathematical Chaos in a deterministic universe and 2) the combinatorial explosion of future events in a deterministic universe. As an aside, we will also consider limitations on deep neural networks as some of the recent successes using those algorithms might suggest they can overcome these shortcomings. They cannot.

Predicting the Future in a Deterministic Chaotic Universe

One of the hallmarks of the Newtonian Age of Science is the success of using deterministic laws in predicting the future. Humanity has discovered many laws and principles of the natural world that have allowed for great advances in transportation, agriculture, medicine, meteorology, and communication. These technologies have significantly affected our success and flourishing as a social species. If a superintelligence has exponentially more cognitive power than a human being, would it be able to make exponential advances in finding and applying deterministic laws and principles in predicting the future?

First, it should be noted that we may not live in a deterministic universe. Current theories in quantum physics suggest that at the subatomic level, events may in fact have a great deal of randomness. However, it is not obvious that the randomness at the subatomic level will necessarily result in randomness at the macro level. Certainly, most of the phenomena we observe at our macro-level do not seem to exhibit a great deal of randomness; rather, there exist reassuring regularities and deterministic behavior. It is plausible that the randomness that occurs at the subatomic level does not have a great impact on our ability to make predictions at the macro level. We will return to this issue later.

A more fundamental concern for making predictions about the future is that, while we may live in a deterministic world, the equations that govern behavior of macroscopic objects exhibit chaos. Chaos Theory studies how seemingly random behavior may result from deterministic equations. Of particular importance to this paper is the fact that equations that exhibit chaos are extremely sensitive to initial conditions. Minute changes in initial values result in vast (and unpredictable) changes in outcomes in the future.

Many of the equations that accurately describe our natural and social world are equations that exhibit chaos. Most physical systems are chaotic (e.g., the gravitational three-body problem, double pendulums, billiards on an oval table, weather phenomena, general relativity, fluid dynamics, heart arrhythmia). Many non-physical systems also exhibit chaotic behavior (e.g., population dynamics, economics, profit models, stock market models). Therefore, any intelligent being will be making predictions about a world that is described by equations that are highly sensitive to initial conditions.

As we will show in the next section, even an exponential increase in the accuracy of initial conditions may only result in a linear increase in prediction time.

The Impact of Chaos on Forecast Horizon

Very simple equations can generate chaotic behavior. Equation 1 presents a simple discrete recurrence relation. The value of a variable x at discrete time, t, depends on its value at previous discrete time, t – 1.

$$x_t = x_{t-1}^2 - 1.9$$

For the purposes of our discussion, let us assume t is defined in discrete seconds. In the graph presented in Figure 2, x_0 is set to 0.5. (Therefore, x_1 is $(0.5)^2 - 1.9$ or -1.65).

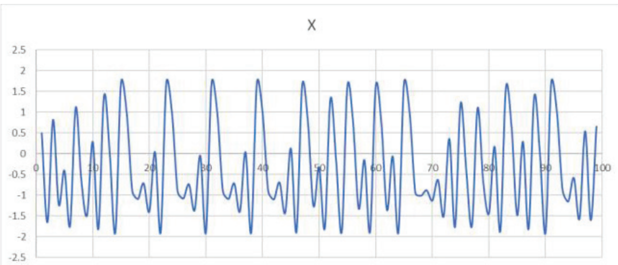

Figure 2 An Illustration of Chaotic Behavior from a Simple Recurrence Relation, $x_t = x^2_{t-1} - 1.9$

As is evident from the graph, this equation describes an oscillating value. The pattern of the oscillation will never repeat. Imagine that x describes the location of some object at a certain time. How accurately can we predict where the object will be at a certain time? If its initial starting position is 0.5, then we can determine its future position with absolute certainty. But suppose we are off in our measurement of its initial position. Suppose we can only measure accurately to one angstrom (one ten-billionth of a meter or 10^{-10} meter). So instead of its starting position being at 0.5, its starting position would be 0.5000000001 m. On Figure 3, we plot both the original curve when x_0 is 0.5, and the new curve where x'_0 is 0.5 plus one angstrom. As can be seen, the curves differ dramatically after around t = 44. In other words, if our original estimate is off by 1 angstrom, we can make accurate predictions up to around 44 seconds before our predictions are no longer accurate (and never regain accuracy).

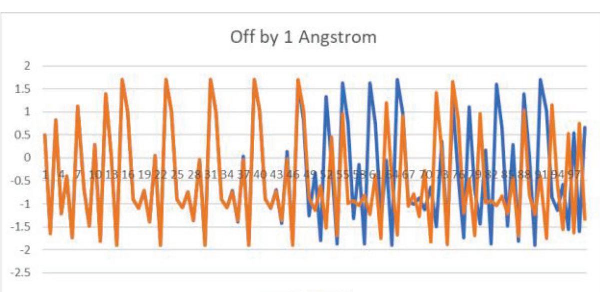

Figure 2 Initial Measurement Is Off By 1 Angstrom

This phenomenon of a forecasting horizon is well known in the field of chaotic mathematics. Small perturbations in initial conditions cause the output of equations to diverge over time. The smaller the perturbation, the longer the two values remain close. The relationship between the magnitude of the forecast horizon varies inversely proportional to the logarithm of the size of the error. The smaller the error, the longer the horizon – but the relationship is exponential. It requires an exponential decrease in the error to result in a linear increase in the forecast horizon.

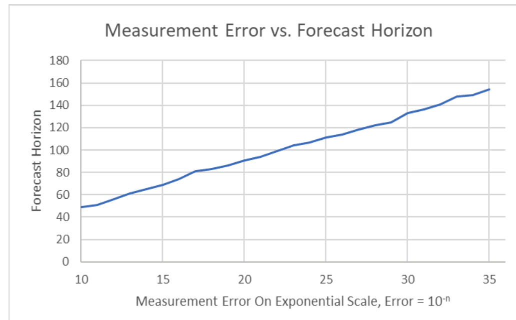

Figure 3 Forecast Horizon By Measurement Error (Note: For example, the error at N=25 is 10^{-25} m)

The implications of this result are astounding when applied to making predictions about the future. Current human technology allows us to make measurements to around 1 angstrom in certain cases. Would a superintelligence be able to make better predictions with increased knowledge? What if the superintelligence had the knowledge to make much finer measurements? In an experiment, the initial measurement was varied by a tiny amount starting with one angstrom. The time was recorded as to when the prediction made by this measurement differed from the actual value by more than 50%. In Figure 4 the measurement error decreases by an order of magnitude along the x-axis. What is extraordinary about this graph is that decreases by several orders of magnitude in the measurement error only result in very modest gains in the forecast horizon. For instance, with a decrease equivalent to 6 orders of magnitude (one million times more accuracy), the forecast horizon goes from 49 seconds to 74 second (1.5 times further). The limits of how far one could look ahead in this simple equation are ultimately capped by the limits in one's measurement ability. In our universe, no measurement is possible below a Planck length, 1.6 x 10^{-35} meters. So even if one were measure at the very limits, over 10^{25} times more accurately than an angstrom, the graph shows that one can only look 3.14 times further into the future.

Some predictions can be made about the value of x. We can predict with 100% accuracy that the value of x will be between -2 and 2. However, we could have experimentally made that estimate even with a fairly inaccurate initial measurement. Increasing the accuracy of the initial measurement has almost no effect on the accuracy of the predicted range of x values.

Similar results can be obtained by looking at other chaotic equations such as those that govern the famous Three-Body Problem. Figure 5 shows a similar experiment for an instance of the Three-Body Problem. Again, notice how an exponential increase in measurement accuracy results in a modest linear increase in forecast horizon. At one angstrom of accuracy, the forecast horizon is at 1,865 seconds; at 10^{-35} accuracy, the forecast horizon is 10,394 – 5.6 times farther.

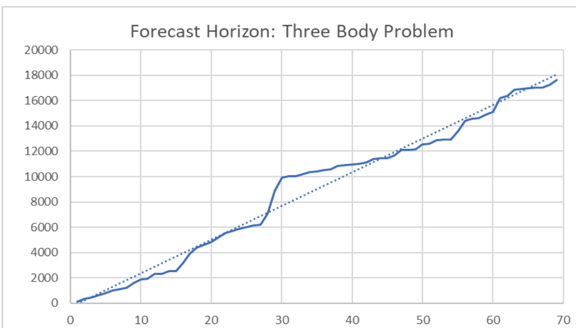

Figure 5 Forecast Horizon for Example Three Body Problem

A conclusion that can be drawn from this analysis is that orders of magnitude gains in accuracy only result in modest improvements in prediction accuracy. The results of such studies are so dismal that researchers in weather forecasting have concluded that local weather forecasting will remain capped in accuracy to 15-16 days regardless of the amount of any additional data or measurement precision – a conclusion that would apply to all intelligent beings, not just humans (Brisch and Kantz, 2019; Zhang et al, 2019).

Predicting the Future with Exponential Branching

Chaotic behavior suggests that increasing computational power will not result in better predictions beyond a modest window. However, even in non-chaotic events, exponential increases in computational power may not result in exponential increases in abilities. This paucity of improvement will be noticeable in any domain where the branching factor is inherently exponential – which is a feature of most future events planning. The canonical example in computer science and artificial intelligence of the impact of exponential explosion is in computer chess.

Several studies have explored how increasing search depth leads to higher chess rankings as indicated in Figure 6. However, ELO rankings are challenging to use to compare vastly unequal players as they are based on winning percentages. Overall winning percentages are a problematic metric to compare computer opponents. Suppose one player beats another player 100% of the time – how much better is that player? Twice as good? One hundred times as good? One million times as good? Instead, we conducted a study that looks at search depth that looks at the chance of finding a better move as the search deepens.

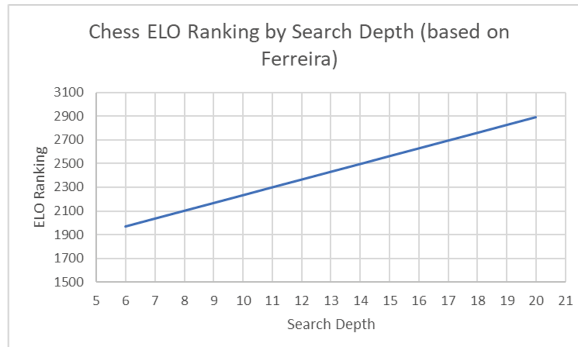

Figure 6 ELO Ranking Estimate Based on Search Depth (from Ferreira, 2013)

In an implementation using the computer chess engine Stockfish, we conducted a study as follows:

- We generated 100 mid-games by having the engine play itself 100 times and stop after 20 moves. The resulting partially completed games all ended up in unique positions.
- For each of the 100 mid-games, the chess engine searched for the best next moves. The engine would be limited in its search by constraining the number of board positions searched in the game tree. These constraints were given in powers of 2, so Level 20 would be a level where the search looked at approximately $2^{20} \sim 1$ million board positions.
- A "Gold Standard" move set was defined as the best three moves as specified by a search that evaluated 2^{34} board positions (which takes approximately 4 hours of clock time).
- We plotted
  1) the percentage of times the best move for a given level was different than the "Gold Standard" (Figure 7), and
  2) the percentage of times the best move for a given level was different than the proceeding level (Figure 8).

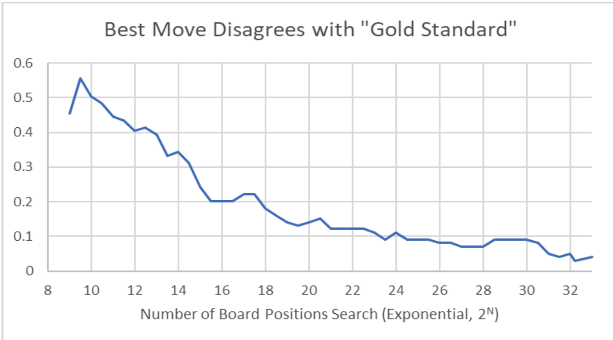

Figure 7 Percentage of the time the best move selected at a level differs from the Gold Standard move

As can be seen from Figure 7, differing from the Gold Standard decreases as the search deepens. The decrease is at first steep, but it begins to level off so that the effect of an exponential increase in the number of board positions search does not result in a significant change in the likelihood of selecting the best move. The algorithm is just as likely to select the best move when searching 2^{23} board positions as it does 2^{30} board positions. As indicated in Figure 8, once a certain threshold of board positions evaluated (around 2^{19} in this chart), the advantage of increasing the search space only results in a 5% chance of selecting a different move at the next level, and that chance is constant across the rest of the graph.

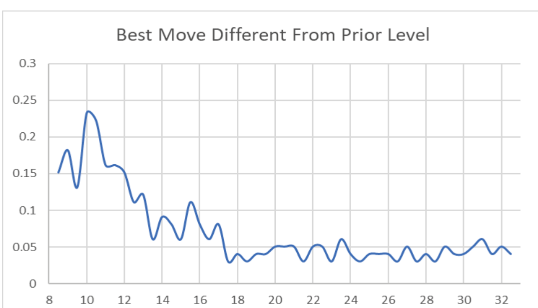

Figure 8 Percentage of Time the Best Move at a Level Differed from the Best Move at the Preceding Level

One interpretation of these results is that an exponential increase in search space follows a law of diminishing returns. Initial gains are quite good – but after a certain point the exponential increase in the search gives only modest gains. This finding corroborates a result of an earlier result by Steenhuisen who looked at search depths up to size 20 (Steenhuisen, 2005).

A Word about (Deep) Neural Networks

Recent successes using deep neural networks might suggest that limitations in search algorithms (such as those discussed in the previous section) can be overcome by using connectionist models. In particular, DeepMind's AlphaZero crushed Stockfish in a 1000-game chess match, seemingly conferring a significant advantage to deep neural networks. In these timed games, AlphaGo dominated. However, neural networks have some shortcomings that may not allow these algorithms to improve at a significant rate even given exponential increases in computational power.

Individual neural networks consistently reach limits in their performance, and no amount of extra data or processing time will improve their performance. Further, adding extra hidden layers, i.e., making the neural network "deep", confers advantages for some problems, but, at some point, adding more layers does not result in better performance.

Most papers describing a neural network experiment comes with a learning curve such as the representative graph given in Figure 9. What this figure indicates is that as the number of training epochs (the amount of training

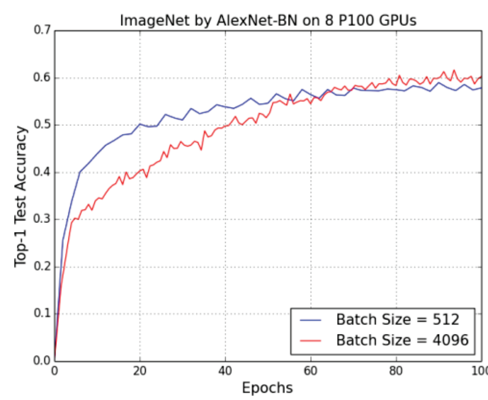

Figure 9 Representative Illustration of the Learning Curve of a Neural Network

increases), the performance of the system gets better. However, as can be clearly seen, after an initial rapid climb, the performance only gets incrementally better, and then it typically has no improvement no matter how much longer it runs or how much data it receives.

Studies also show that increasing the depth of a neural network by adding hidden layers may result in some initial

method	# layers	# params	error (%)
Maxout [10]			9.38
NIN [25]			8.81
DSN [24]			8.22
FitNet [35]	19	2.5M	8.39
Highway [42, 43]	19	2.3M	7.54 (7.72±0.16)
Highway [42, 43]	32	1.25M	8.80
ResNet	20	0.27M	8.75
ResNet	32	0.46M	7.51
ResNet	44	0.66M	7.17
ResNet	56	0.85M	6.97
ResNet	110	1.7M	**6.43** (6.61±0.16)
ResNet	1202	19.4M	7.93

Figure 10 Increasing the Number Hidden Layers and Its Effect on Error (He et al, 2016)

gains, but additional layers may actually result in worse performance. A seminal study on image processing using ResNet found that increasing the number of hidden layers found diminishing returns as indicated in the chart given in Figure 10.

Why We Still Might Want to Be Concerned About Superintelligences

This paper attempts to dispel the fears that machines in the future will be “billions” of times more intelligent than humans. Even with massive improvements in computational power, such machines only have the potential for modest linear improvements in intelligence. Nonetheless, linear improvements are still improvements. A machine that can predict stock market behavior five seconds ahead of another can dominate trading, the basis of high-frequency trading (HFT) algorithms. Making a better chess move than your opponent only 5% of the time (and never making a worse move) will result in a 60-70% winning percentage. When we look to a future with superintelligent machines, we do need to be concerned that machines will supplant many workers, not just because they are cheaper, but because they are better. An accountant that can find tax savings that are 3% better is going to be preferred. A machine won’t have to be a 1000 times smarter to take your job – just 3% better. Nick Bostrom’s concern that autonomous machines will outsmart us, manipulate us and run circles around us seems unlikely given that they will only be modestly better at predicting and planning the future, and we can still build non-autonomous machines that can be used as tools to match wits with those autonomous machines.

Conclusions

Machines with extraordinary computational power will be built in our future. Their ability to perform computations will certainly exceed the human body’s ability to perform computations. As those systems improve, their memory capacity and processor speed may outstrip human capacity by many orders of magnitude. This paper argues that those machines will not outstrip human intelligence by many orders of magnitude. In fact, their level of intelligence, as measured by their ability to predict and plan for the future, may only be marginally better than human ability. The limitations mentioned in this paper are two-fold: 1) limitations based on the inability to predict future events that are based on chaotic equations, and 2) limitations based on diminishing returns of analyzing exponentially expanding search spaces. Chaos is of particular concern in predicting the future, even in a deterministic world. We close this paper by noting that quantum physics suggest that the physical world is not deterministic – randomness exists at the subatomic level. And, because chaotic equations are so sensitive to initial values, a fluctuation at the quantum level, even at the dimensions of a Planck length (10^{-35} m), can very quickly propagate through the system resulting in changes at the macro-scale. Thus, any efforts for accurate predictions of future events on the human scale beyond a few minutes, hours or days, may be theoretically impossible. No matter how smart you are.

References

Bostrom, N. (2014) Superintelligence: Paths, Dangers, Strategies (1st ed.). Oxford University Press, Inc., New York, NY, USA.

Brisch, J. and Kantz., (2019) Power law error growth in multi-hierarchical chaotic systems -- a dynamical mechanism for finite prediction horizon in weather forecasts, *New J. of Physics*, V. 21.

CNBC, (2018), AI will be ‘billions of times’ smarter than humans and man needs to merge with it, expert says, https://www.cnbc.com/2018/02/13/a-i-will-be-billions-of-times-smarter-than-humans-man-and-machine-need-to-merge.html

Ferreira, Diogo R., (2013) The Impact of the Search Depth on Chess Playing Strength, ICGA Journal, vol. 36, no. 2, pp. 67-80.

Good, I. J. (1965) Speculations Concerning the First Ultraintelligent Machine, Advances in Computers, vol. 6, 1965.

Hommel B. (2015) The theory of event coding (TEC) as embodied-cognition framework. *Frontiers in psychology*, 6, 1318.

Kurzweil, R. (1999) The Age of Spiritual Machines: When Computers Exceed Human Intelligence. New York: Viking, 1999.

He, K. Zhang, X., Ren, S. , Sun, J, (2016) Deep Residual Learning for Image Recognition,The IEEE Conference on Computer Vision and Pa ttern Recognition (CVPR), June, 2016.

Steenhuisen, J. R. (2005). New Results in Deep-Search Behaviour. ICGA Journal, Vol. 28, No. 4, pp. 203–213.

Vinge, V. (1993) The coming technological singularity: How to survive in the post-human era, Whole Earth Review.

You, Y., Zhang, Z, Hsieh, C-J, Demmel, J, Keutzer, K.(2018). ImageNet Training in Minutes. Proceedings of the 47th International Conference on Parallel Processing (ICPP 2018). ACM, New York, NY, USA.

Zhang, F., Qiang Sun, Y., Magnusson, L., Buizza, R., Lin, S.-J., Chen, J.-H., and Emanuel, K. (2019). What is the Predictability Limit of Midlatitude Weather? *J. Atmos. Sci.*