

APPLICATION

GFLOW: software for modelling circuit theory-based connectivity at any scale

Paul B. Leonard^{1*}, Edward B. Duffy², Robert F. Baldwin¹, Brad H. McRae³, Viral B. Shah⁴ and Tanmay K. Mohapatra⁴

¹Department of Forestry and Environmental Conservation, Clemson University, Clemson, SC 29634, USA;

²Cyberinfrastructure Technology Integration Group, Clemson University, Clemson, SC 29634, USA; ³The Nature Conservancy, Fort Collins, CO 80523, USA; and ⁴Julia Computing, Inc., Bengaluru 560034, India

Summary

1. Increasing habitat connectivity is important for mitigating the effects of climate change, landscape fragmentation and habitat loss for biodiversity conservation. However, modelling connectivity at the relevant scales over which these threats occur has been limited by computational requirements.
2. Here, we introduce the open-source software GFLOW, which massively parallelizes the computation of circuit theory-based connectivity. The software is developed for high-performance computing, but scales to consumer-grade desktop computers running modern Linux or Mac OS X operating systems.
3. We report high computational efficiency representing a 173× speedup over existing software using high-performance computing and a 8.4× speedup using a desktop computer while drastically reducing memory requirements.
4. GFLOW allows large-extent and high-resolution connectivity problems to be calculated over many iterations and at multiple scales. We envision GFLOW being immediately useful for large-landscape efforts, including climate-driven animal range shifts, multitaxa connectivity, and for the many developing use-cases of circuit theory-based connectivity.

Key-words: circuit theory, gene flow, landscape connectivity, landscape ecology

Introduction

As ecologists attempt to model increasingly complex systems, computational limits have become a factor that influence experimental design, analysis and inference (Schadt *et al.* 2010; Beckage, Gross & Kauffman 2011; Visser *et al.* 2015). A case in point is resistance-based methods for estimating habitat connectivity, which encounter computing limits as more complex problems are modelled and expectations for scale-based solutions increase (Moilanen 2011). Yet improving connectivity is the most frequently cited strategy to mitigate ongoing threats of climate change, wildlife habitat loss, fragmentation and/or degradation for biodiversity conservation (Crooks & Sanjayan 2006; Heller & Zavaleta 2009; Nunez *et al.* 2013). Connectivity is broadly defined as the degree to which a landscape facilitates movement of organisms between resource patches (Taylor *et al.* 1993). More specifically, it is described by two primary components: (i) the structural or physical component and (ii) the functional or behavioural component (Tischendorf & Fahrig 2000). Methods used to capture these components have expanded from least-cost path analyses identifying narrow swaths of land to complex functional metrics,

both empirical and theoretical, that emphasize matrix permeability over large areas (Kindlmann & Burel 2008). Recently, there has been a proliferation of metrics that account for gradients of connectivity across multiple networks of patches in the landscape (Wagner & Fortin 2005; McRae 2006; Saura & Pascual-Hortal 2007; Theobald *et al.* 2012). These methods are largely based on a graph-theoretic framework described by Urban & Keitt (2001) and attempt to incorporate multiple components of connectivity by estimating resistance (i.e. friction) to movement between patches.

Circuit theory-based applications extend previous graph theory approaches to incorporate functional connectivity. They treat landscapes as conductive surfaces, using electrons flowing across an analogous circuit to model individuals or genes moving across the landscape (McRae *et al.* 2008). Areas with higher resistance to movement or gene flow are assigned higher resistance in the analogous circuit. This analogy is useful because it facilitates the prediction of multiple pathways accounting for both structure and function by having mathematical relationships with random walk theory (Doyle & Snell 1984). These underpinnings provide circuit theory a direct link to landscape and population genetics and movement ecology (McRae 2006; McRae *et al.* 2008).

Many studies have demonstrated the relationship between circuit theory and intraspecies long-term movement, including

*Correspondence author. E-mail: pbleona@clemson.edu

gene flow (Shirk *et al.* 2010; Moore *et al.* 2011; Koen *et al.* 2012). Complexity of the modelling approach is increasing, as there has been recent expansion to describe functional connectivity for multiple taxa (Walpole *et al.* 2012; Koen *et al.* 2014; Pelletier *et al.* 2014; Leonard *et al.* 2016) but the broad availability of spatial data across large extents at fine grains has outpaced our ability to process, manage and incorporate these data into connectivity models (Rubio *et al.* 2015). The methods sections of numerous studies report that raster data resolutions were coarsened in order to model resistance-based connectivity over large extents (e.g. Schwartz *et al.* 2009; Shirk *et al.* 2010; Lawler *et al.* 2013; Castillo *et al.* 2016). Because the goal of large-extent connectivity mapping is often incorporated with efforts to identify areas of conservation priority (Beier *et al.* 2011), and these may be scale- and species-dependent, it is important to create a process by which connectivity can be assessed at multiple spatial resolutions over many iterations, that is a methodology robust to the size of the ecological network or problem.

In this paper, we present new software that greatly expands the capacities of CIRCUITSCAPE, the most used and cited circuit theory application. In 2015 alone, it appeared in 80 peer-reviewed journal articles (a 40% increase from 2014) and has been widely implemented by conservation organizations around the world. Currently, CIRCUITSCAPE relies on standard linear solvers (mathematical software libraries) that limit authors' ability to solve large problems and often lead to analytical trade-offs. Modelling decisions 'forced' by computational limits may inhibit ability to capture fine and broad-scale ecological processes, and represent them in applied contexts. Here, we describe a freely available, open-source, software program GFLOW, which massively parallelizes the computation of circuit theory allowing for simultaneous large-extent and fine-grained analyses of connectivity. Our goal is to provide ecologists, biologists and conservation planners the ability to reliably calculate gene flow, habitat and landscape connectivity for virtually any sized problem using high-performance compute clusters, scalable cloud computing resources, and for moderate-sized problems using desktop computing. Below, we describe the general approach of GFLOW and detail potential applications and future improvements.

Overview of approach

IMPLEMENTATION DETAILS

We developed GFLOW using the C programming language on version 6.7 of the Scientific Linux operating system distribution. Parallelization of GFLOW was made possible using OpenMPI (www.open-mpi.org), an open-source implementation of the Message Passing Interface (MPI). We leveraged the Portable, Extensible Toolkit for Scientific Computation (PETSc) (www.mcs.anl.gov/petsc/) in conjunction with the Parallel High Performance Preconditioners (hypre) library (<http://computation.llnl.gov/projects/hypre-scalable-linear-solvers-multigrid-methods>) to manage our data structures and

perform the numerical analysis. The linear system is solved using PETSc's conjugate gradient method. Within the hypre library, we took advantage of the BoomerAMG preconditioner, a parallel implementation of the Ruge–Stüben algebraic multigrid algorithm (See Fig. S1 for workflow and Appendix S1 for installation procedure, Supporting Information). The numerical and combinatorial methods are described in detail by Shah (2007).

DATA REQUIREMENTS, INPUT AND OUTPUT

Circuit theory-based connectivity is modelled by assigning each pixel in a landscape matrix a resistance value indicating the degree of landscape permeability for the flow of electrons (e.g. animals, genes). Resistance values may be empirically derived estimates based on movement or genetic data or derived via expert opinion (Spear *et al.* 2010; Zeller, McGarigal & Whiteley 2012). GFLOW accepts the resistance surface stored as an ASCII grid file, with each grid cell being converted to a node in the corresponding circuit. The most common type of analysis includes connecting source and destination locations (i.e. 'focal nodes') in a pairwise fashion that may represent habitat patches, populations or other areas to connect (e.g. protected areas). Current densities in intervening grid cells are measured as current travels through the resistance surface between source/destination nodes. These locations are also accepted as an ASCII grid file or as a tab-delineated text file with node ID, and x,y locations relative to the input resistance surface. If an ASCII file is input, the program will create the latter input automatically for computational efficiency. Options for output are limited to a summation of per-cell current density (in amperes) for all source/destination pairwise node computations, for a random selection of pairwise nodes or for each individual pairwise computation in addition to the summation. These values are reported back in ASCII grid files for easy import into a geographic information system. Additional outputs include source/destination node effective resistances, which are most often used in landscape genetics studies, and a convergence factor discussed below.

ADDITIONAL FUNCTIONALITY

In addition to calculating current density between a set of pre-determined nodes, GFLOW offers a novel metric for those wishing to calculate landscape connectivity between random node locations. Koen *et al.* (2014) described methods to reduce node location bias by including a buffer around their study area encasing randomly placed nodes. These methods allow practitioners to conduct fewer pairwise computations or to calculate connectivity where little information is known about relative node quality or exact node location. However, it remains unclear how many random node connections are required for any given problem to fully saturate the landscape with current and thus converge upon a solution. To address this concern, we provide a convergence factor (correlation function), which is calculated with each pairwise solve (i.e. iteration) and ranges from 0 to 1:

$$C_{(ij)} = \text{corr}(X_{(i)}, X_{(j)}),$$

where $X_{(i)}$ and $X_{(j)}$ are consecutive random pairwise summations of current density at iterations i and j . With each iteration, the delta convergence approaches a value of zero, indicating additional connections very marginally affect the global simulation. The user can set a stopping criterion based on a class of nines most often seen in reliability studies (Piedad & Hawkins 2001). For example, if an additional pairwise calculation and subsequent summation only differs from the existing summation by 0.0001, then $C = 0.9999$ (i.e. 'four nines') and the program can cease calculations and create output.

VISUALIZATION

Because basic file management and display of large ASCII text files is cumbersome, we created a stand-alone visualization tool (GVIEW) that converts ASCII files to various image formats for rapid inspection. This tool can be particularly useful for visualizing intermediate output or individual pairwise calculations that do not require a GIS (Fig. 1). GVIEW relies on the image-processing library IMAGEMAGICK. The GVIEW default for scaling current density values is based on exponential classification but allows the user to choose among manual classification or statistical breaks. In addition, the user may choose between pre-packaged colour ramps or create custom colour schemes.

Example use

Because landscape connectivity is a dynamic concept, and is often species and scale specific, it may be best characterized in the context of environmental change (Crooks & Sanjayan 2006). These considerations add to the complexity of connectivity analyses that are already computationally challenging. Through many iterations and scenarios, GFLOW enables in-depth exploration of these dynamics and their

interactions using sensitivity and uncertainty analyses that have been previously impracticable. With a developmental version of the software, Leonard *et al.* (2016) created multi-species connectivity maps across various spatial and temporal scales to estimate changing landscape connectivity patterns in the south-eastern United States (Fig. 2). The authors' ability to explore the effects of increasing urbanization and sea-level rise using species-specific models with multiple resistances produced the first large-landscape forecast of its kind. In addition, this GFLOW example allows conservation practitioners with a species-specific mandate to manage for enduring features (e.g. riparian and river corridor connectors) that provide broader value by increasing landscape permeability.

Software performance results

SUPERCOMPUTER DEPLOYMENT

For performance testing, we used a moderately sized circuit theory-based connectivity problem at 270-m spatial resolution across 588 predicted core-habitat locations covering 2.5 million km² of the south-eastern United States. We chose this spatial resolution in order to make comparisons with existing software possible while demonstrating potential utility for large-landscape conservation planning. The full pairwise map consists of more than 172.5 thousand pairwise calculations requiring the calculation of more than 34.25 million current density measurements at each iteration. We deployed GFLOW on the Palmetto Cluster at Clemson University, which is the sixth fastest academic supercomputer in the United States (top500.org, 2016). This deployment nearly achieved a linear or 'ideal' speedup ratio (i.e. where doubling the number of processors halves the compute time) with a mean parallel efficiency of 0.93 across our multiple tests (Fig. 3). Efficiency is used to describe the scalability of a computing problem (typically ranging from 0 to 1), and we measure it where:

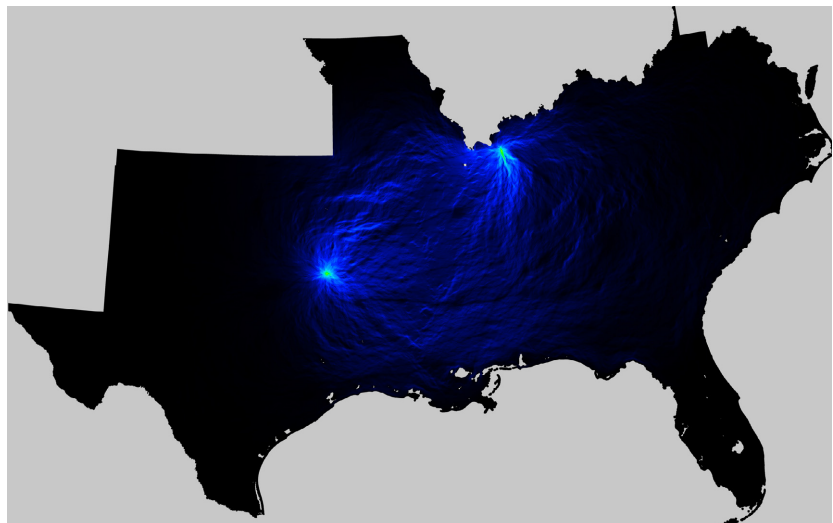


Fig. 1. Example image from the stand-alone viewer GVIEW with default settings for one pairwise calculation from our performance test model. The geographic region encompasses the south-eastern United States including Texas (2.5 million km² or 34.25 million pixels), which is represented by a large text file.

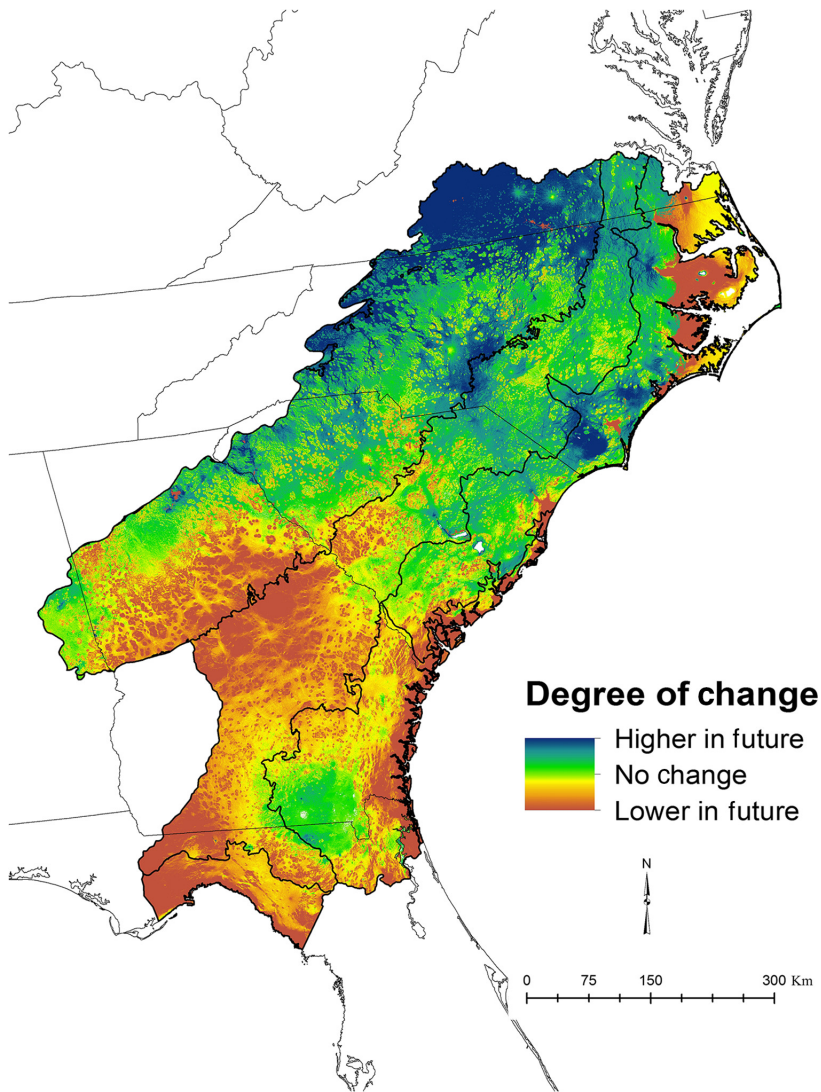


Fig. 2. Circuit theory-based connectivity analysis using GFLOW to predict changes in landscape connectivity patterns into the year 2100 based on a composite of multiscaled habitat connectivity models for six individual species in the south-eastern United States. Degree of change is represented by the relative distribution of values with the assumption that future values, although lower than contemporary values, will still be important to dispersing animals. Adapted from Leonard *et al.* (2016).

$$E = \frac{T_1}{pT_p},$$

T_1 = processing time of sequential execution, T_p = processing time in parallel execution, p = the number of processors.

DESKTOP DEPLOYMENT

We developed GFLOW for use in a high-performance computing environment to improve the capacity of ecologists and conservation planners to capture complexity related to flow at multiple scales. However, we appreciate that access to supercomputers is limited and that a typical ecologist may welcome a version of the software that can be executed on a desktop computer, especially because existing software presents challenges for desktop computation when data sets are large, as is often the case in regional planning applications (Leonard *et al.* 2014). To demonstrate the potential of GFLOW on a consumer-grade desktop computer, we compiled a version for Apple, Inc.'s OS X operating systems and executed the same inputs as those reported in the supercomputer deployment. Execution on the desktop computer was relatively fast

(Table 1). However, larger problems requiring many thousands of iterations will remain time-intensive as parallel efficiency was considerably lower ($E = 0.81$) than our supercomputing tests and thus relative gains in speed (i.e. 'speedup') quickly departed from 'ideal' (Fig. 3).

Comparison with existing software

There are at least three software packages that are in use by ecologists to calculate circuit theory connectivity metrics: CIRCUITScape (Shah & McRae 2008), GRAPHAB (Foltête, Clauzel & Vuidel 2012) and the R package GDISTANCE (Van Etten 2012). We chose to focus our comparisons on the most widely used, and to the best of our understanding, most computationally efficient existing application for calculating current density – CIRCUITScape. Rather than focus solely on gains in speed, we chose to demonstrate the decreased memory required for GFLOW to solve large problems. The version of CIRCUITScape we tested (v4.0.5) does have a task-level parallel option for computation but due to its high memory demands (30 GB per process), even a supercomputer could only use four CPUs at any given time to solve the problem. GFLOW has much lower

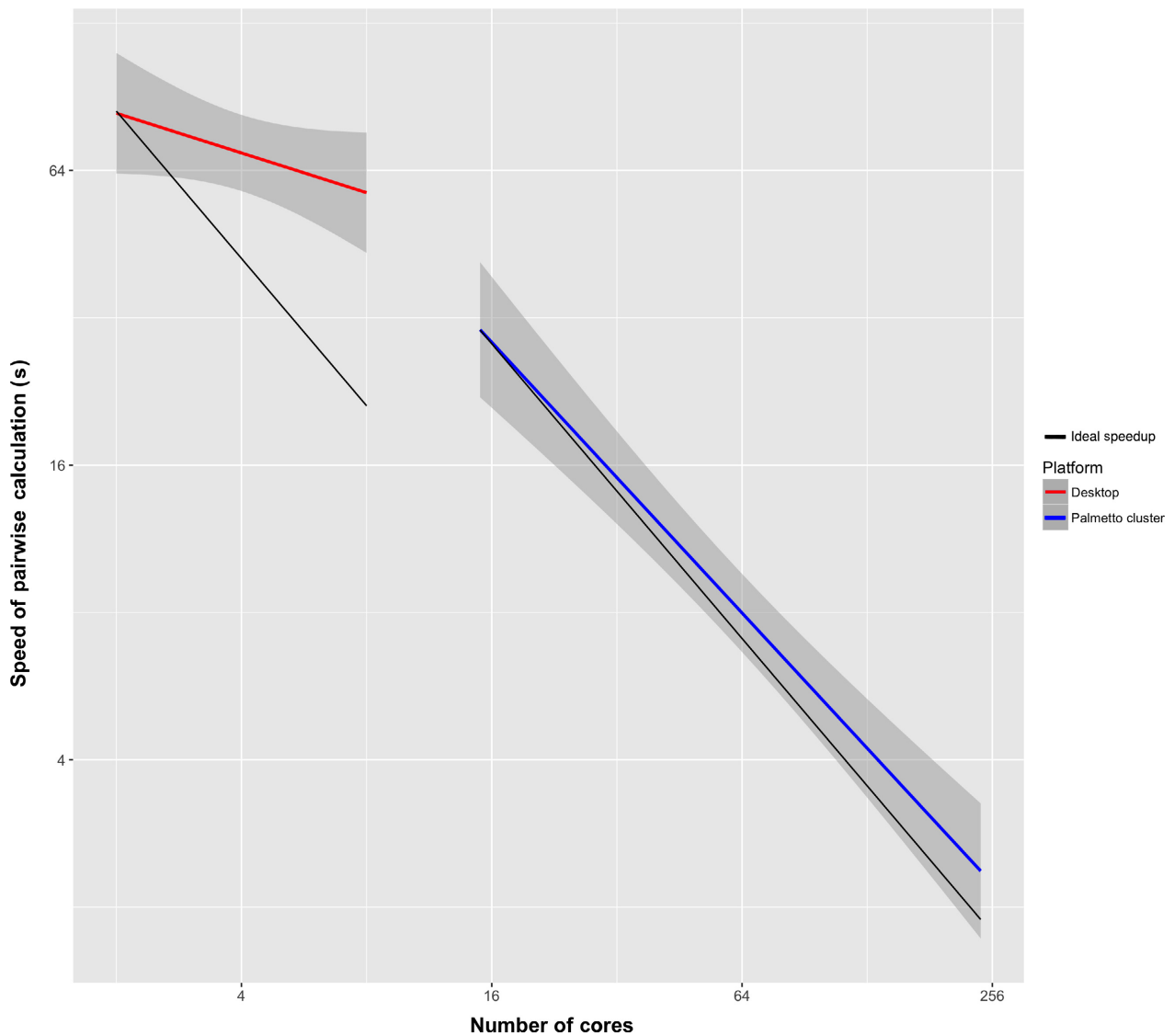


Fig. 3. Performance of GFLOW on the Palmetto Cluster supercomputer compared with a desktop computer equipped with eight cores. On the supercomputer, the speedup and efficiency are nearly 'ideal', which falls within the standard error of our measurements. However, the desktop computer quickly diverges from a linear speedup. We tested GFLOW with a minimum of 15 cores on the supercomputer and two cores on the desktop.

Table 1. Timing and resource usage for GFLOW as deployed on the supercomputer *Palmetto Cluster* and a desktop computer. Required memory, initialization and mean pairwise solve time for 100 iterations of the test problem are reported. The same desktop computer was used for testing CIRCUITSCAPE and GFLOW

Palmetto cluster				Desktop			
# of CPUs	RAM per CPU (GB)	Initialization (s)	Pairwise speed (s)	# of CPUs	RAM per CPU (GB)	Initialization (s)	Pairwise speed (s)
GFLOW							
240	1	18	2.6	8	2	13	58.2
120	1	17	4.2	4	4	10	68.1
60	1	17	7.3	2	8	10	84.6
30	1	16	18.1				
15	1	15	30.2				
CIRCUITSCAPE							
1	30	576	450.1	1	30	659	488.6

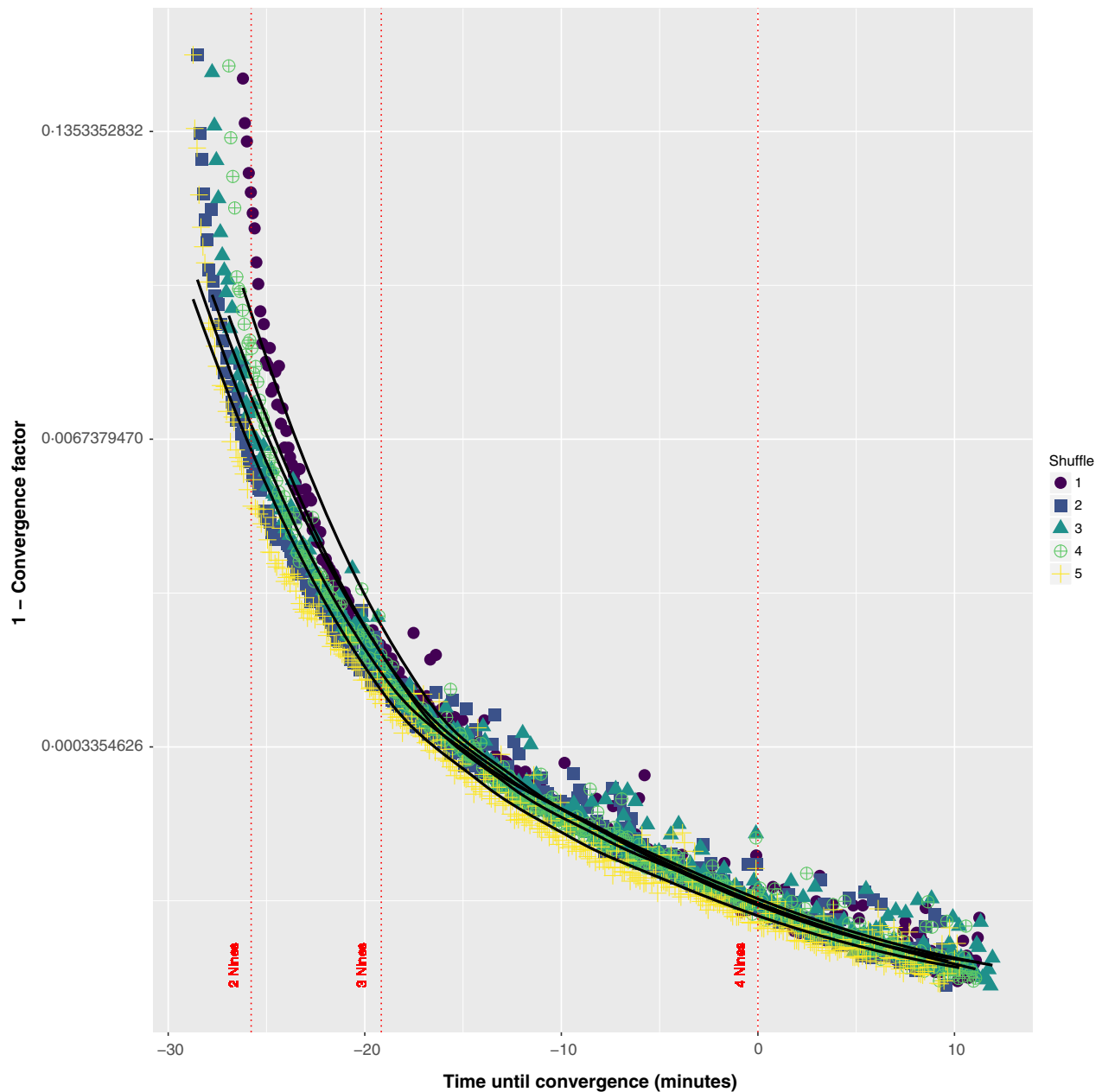


Fig. 4. Five random shuffles ($n = 400$ pairwise solves) from all possible point pairs ($n = 172\,578$) describing the amount of wall time GFLOW required to converge upon the solution (using 60 CPUs) based on our stopping criteria ($C \geq 0.9999$ or four ‘nines’). Stopping criteria is an option in GFLOW, and we encourage users to explore similar convergence functions based on the resources available to solve big problems.

memory demands and executed the test model easily resulting in a $173\times$ pairwise solve speedup using the Palmetto super-computer with 240 cores, or an $8.4\times$ solve speedup using a desktop computer with eight cores (Table 1). In comparison, we estimated that it would take a sequential solver (e.g. CIRCUITScape) 2.6 years to calculate the test problem solution we calculated in 5 days with the Palmetto Cluster.

Discussion/Future work

GFLOW is a timely advance and may have broad application in ecology, as there has been recent expansion of the use of circuit theory to predict the spread of infectious disease and invasive

species (Andraca-Gómez *et al.* 2015; Cowley, Johnson & Pocock 2015), wildfires (Gray & Dickson 2015) and connectivity in the face of climate change (Lawler *et al.* 2013; Razgour 2015). In all of these instances, authors were forced to coarsen their analyses and/or create unique analytical steps to make computation possible, including discretizing the landscape into tiles (see Pelletier *et al.* 2014) and distance-based thresholds or a combination of similar methods (see Roever, van Aarde & Leggett 2013; Castillo *et al.* 2016). While these computational compromises may be acceptable in some cases, it is not difficult to imagine the expansion of circuit theory problems that require many thousands of node locations over large areas at fine spatial grains as remote sensing technology continues to

produce higher-quality data. In other words, we increasingly have the ability to observe problems at the scales in which we need to model them. GFLOW fundamentally changes the limiting factor in calculating circuit theory-based connectivity from compute resources to our conceptual understanding of the system being modelled via inputs and outputs. For many small and moderately sized problems, desktop computation with GFLOW will greatly decrease computational demands over existing software but it can also be deployed in a dynamic, cloud cluster computing environment that a user may request on demand (e.g. Amazon Elastic Compute Cloud).

Future development of GFLOW will focus on increased functionality by handling additional data formats for both input and output. We are also exploring additional modes for testing the effects of previous computational trade-offs and for output options (e.g. effective resistance only for landscape genetics studies). Finally, we hope to offer a publically accessible cloud implementation in future. We currently have no intention of developing GFLOW for a windows-based operating system because they make up only 0.2% of supercomputers in widespread use (top500.org, 2016). Likewise, desktop deployment is limited to Linux and Mac operating systems because of scalability (the ability to maximize efficiency of hardware), flexibility and customization and ease of use for cross-platform Unix-like systems. For windows users without access to the above alternatives, we recommend virtualizing a Linux operating system distribution, but performance will suffer.

Closing comments

GFLOW is designed to solve large connectivity problems and to allow researchers to investigate a variety of ecological phenomena at multiple scales using multiple iterations. GFLOW has solved problems almost $9\times$ larger, or more than 300 million current density calculations per iteration, than the example provided above, using the Palmetto Cluster supercomputer. However, for such large-scale problems, we encourage users to explore the random pairwise computations option using a convergence factor. We tested a stopping criterion of four nines for random pairwise connections of the test model and achieved a global correlation with all pairs of 0.97 using only 231 calculations taking <30 min (Fig. 4). We envision GFLOW being immediately useful for large-landscape efforts, including climate-driven animal range shifts (Lawler *et al.* 2013), multi-taxa connectivity (Leonard *et al.* 2016), and for the many developing use-cases of circuit theory-based connectivity.

Authors' contributions

P.L. and E.D. conceived GFLOW and its features; E.D. led software development and testing; P.L. and R.B. conceived the data analysis, data acquisition and interpretation; P.L. led the writing of the manuscript; B.M., V.S. and T.M. conceived of and provided open-source access to the program CIRCUITSCAPE.

Acknowledgements

We would like to acknowledge the Cyberinfrastructure Technology Integration group at Clemson University for supercomputer and research support. We thank

Ron Sutherland and the Wildlands Network for helpful discussions and for providing data to explore circuit theory modelling limitations. We also acknowledge the Department of Forestry and Environmental Conservation at Clemson University.

Data accessibility

The GFLOW and GVIEW software packages, as well as the example data presented in the manuscript, are available on github here: <https://github.com/Pbleonard/gflow/releases/latest> with the following DOI: <http://doi.org/10.5281/zenodo.60726>.

References

- Andraca-Gómez, G., Ordano, M., Boege, K., Domínguez, C.A., Piñero, D., Pérez-Ishiwara, R., Pérez-Camacho, J., Cañizares, M. & Fornoni, J. (2015) A potential invasion route of *Cactoblastis cactorum* within the Caribbean region matches historical hurricane trajectories. *Biological Invasions*, **17**, 1397–1406.
- Beckage, B., Gross, L.J. & Kauffman, S. (2011) The limits to prediction in ecological systems. *Ecosphere*, **2**, 1–2.
- Beier, P., Spencer, W., Baldwin, R.F. & McRae, B.H. (2011) Toward best practices for developing regional connectivity maps. *Conservation Biology*, **25**, 879–892.
- Castillo, J.A., Epps, C.W., Jeffress, M.R., Ray, C., Rodhouse, T.J. & Schwalms, D. (2016) Replicated landscape genetic and network analyses reveal wide variation in functional connectivity for American pikas. *Ecological Applications*, **26**, 1660–1676.
- Cowley, D.J., Johnson, O. & Pocock, M.J.O. (2015) Using electric network theory to model the spread of oak processionary moth, *Thaumetopoea processionea*, in urban woodland patches. *Landscape Ecology*, **30**, 905–918.
- Crooks, K.R. & Sanjayan, M.A. (eds) (2006) *Connectivity Conservation*. Cambridge University Press, Cambridge, UK.
- Doyle, P.G. & Snell, J.L. (1984) Random walks and electric networks. *AMC*, **10**, 12.
- Foltête, J.C., Clauzel, C. & Vuidel, G. (2012) A software tool dedicated to the modelling of landscape networks. *Environmental Modelling & Software*, **38**, 316–327.
- Gray, M.E. & Dickson, B.G. (2015) A new model of landscape-scale fire connectivity applied to resource and fire management in the Sonoran Desert, USA. *Ecological Applications*, **25**, 1099–1113.
- Heller, N.E. & Zavaleta, E.S. (2009) Biodiversity management in the face of climate change: a review of 22 years of recommendations. *Biological Conservation*, **142**, 14–32.
- Kindlmann, P. & Burel, F. (2008) Connectivity measures: a review. *Landscape Ecology*, **23**, 879–890.
- Koen, E.L., Bowman, J., Garraway, C.J., Mills, S.C. & Wilson, P.J. (2012) Landscape resistance and American marten gene flow. *Landscape Ecology*, **27**, 29–43.
- Koen, E.L., Bowman, J., Sadowski, C. & Walpole, A.A. (2014) Landscape connectivity for wildlife: development and validation of multispecies linkage maps. *Methods in Ecology and Evolution*, **5**, 626–633.
- Lawler, J.J., Ruesch, A.S., Olden, J.D. & McRae, B.H. (2013) Projected climate-driven faunal movement routes. *Ecology Letters*, **16**, 1014–1022.
- Leonard, P.B., Baldwin, R.F., Duffy, E.B., Lipscomb, D.J. & Rose, A.M. (2014) High-throughput computing provides substantial time savings for landscape and conservation planning. *Landscape and Urban Planning*, **125**, 156–165.
- Leonard, P.B., Sutherland, R., Baldwin, R.F., Fedak, D., Carnes, R. & Montgomery, A. (2016) Landscape connectivity losses due to sea level rise and land use change. *Animal Conservation*. doi:10.1111/acv.12289.
- McRae, B.H. (2006) Isolation by resistance. *Evolution*, **60**, 1551–1561.
- McRae, B.H., Dickson, B.G., Keitt, T.H. & Shah, V.B. (2008) Using circuit theory to model connectivity in ecology, evolution and conservation. *Ecology*, **10**, 2712–2724.
- Moilanen, A. (2011) On the limitations of graph-theoretic connectivity in spatial ecology and conservation. *Journal of Applied Ecology*, **48**, 1543–1547.
- Moore, J.A., Tallmon, D.A., Nielsen, J. & Pyare, S. (2011) Effects of the landscape on boreal toad gene flow: does the pattern–process relationship hold true across distinct landscapes at the northern range margin? *Molecular Ecology*, **20**, 4858–4869.
- Nunez, T.A., Lawler, J.J., McRae, B.H., Pierce, D.J., Krosby, M.B., Kavanagh, D.M., Singleton, P.H. & Tewksbury, J.J. (2013) Connectivity planning to address climate change. *Conservation Biology*, **27**, 407–416.

- Pelletier, D., Clark, M., Anderson, M.G., Rayfield, B., Wulder, M.A. & Cardille, J.A. (2014) Applying circuit theory for corridor expansion and management at regional scales: tiling, pinch points, and omnidirectional connectivity. *PLoS One*, **9**, e84135.
- Piedad, F. & Hawkins, M. (2001) *High Availability: Design, Techniques, and Processes*. Prentice Hall Professional, New York, USA.
- Razgour, O. (2015) Beyond species distribution modeling: a landscape genetics approach to investigating range shifts under future climate change. *Ecological Informatics*, **30**, 250–256.
- Roever, C.L., van Aarde, R.J. & Leggett, K. (2013) Functional connectivity within conservation networks: delineating corridors for African elephants. *Biological Conservation*, **157**, 128–135.
- Rubio, L., Bodin, Ö., Brotons, L. & Saura, S. (2015) Connectivity conservation priorities for individual patches evaluated in the present landscape: how durable and effective are they in the long term? *Ecography*, **38**, 782–791.
- Saura, S. & Pascual-Hortal, L. (2007) A new habitat availability index to integrate connectivity in landscape conservation planning: comparison with existing indices and application to a case study. *Landscape and Urban Planning*, **83**, 91–103.
- Schadt, E.E., Linderman, M.D., Sorenson, J., Lee, L. & Nolan, G.P. (2010) Computational solutions to large-scale data management and analysis. *Nature Reviews. Genetics*, **11**, 647–657.
- Schwartz, M.K., Copeland, J.P., Anderson, N.J., Squires, J.R., Inman, R.M., McKelvey, K.S., Pilgrim, K.L., Waits, L.P. & Cushman, S.A. (2009) Wolverine gene flow across a narrow climatic niche. *Ecology*, **90**, 3222–3232.
- Shah, V.B. (2007) *An interactive system for combinatorial scientific computing with an emphasis on programmer productivity*. PhD thesis, University of California, Santa Barbara, CA, USA.
- Shah, V.B. & McRae, B.H. (2008) Circuitscape: a tool for landscape ecology. pp. 62–66 in *Proceedings of the 7th Python in Science Conference*.
- Shirk, A.J., Wallin, D.O., Cushman, S.A., Rice, C.G. & Warheit, K.I. (2010) Inferring landscape effects on gene flow: a new model selection framework. *Molecular Ecology*, **19**, 3603–3619.
- Spear, S.F., Balkenhol, N., Fortin, M.J., McRae, B.H. & Scribner, K.I.M. (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular Ecology*, **19**, 3576–3591.
- Taylor, P.D., Fahrig, L., Henein, K. & Merriam, G. (1993) Connectivity is a vital element of landscape structure. *Oikos*, **68**, 571–573.
- Theobald, D.M., Reed, S.E., Fields, K. & Soule, M. (2012) Connecting natural landscapes using a landscape permeability model to prioritize conservation activities in the United States. *Conservation Letters*, **5**, 123–133.
- Tischendorf, L. & Fahrig, L. (2000) How should we measure landscape connectivity? *Landscape Ecology*, **15**, 633–641.
- Urban, D. & Keitt, T. (2001) Landscape connectivity: a graph-theoretic perspective. *Ecology*, **82**, 1205–1218.
- Van Etten, J. (2012) *Gdistance: Distances and Routes on Geographical Grids*. R package version 1.1-4. Available at: CRAN.R-project.org/package=gdistance.
- Visser, M.D., McMahon, S.M., Merow, C., Dixon, P.M., Record, S. & Jongejans, E. (2015) Speeding up ecological and evolutionary computations in R: essentials of high performance computing for biologists. *PLoS Computational Biology*, **11**, e1004140.
- Wagner, H.H. & Fortin, M.J. (2005) Spatial analysis of landscapes: concepts and statistics. *Ecology*, **86**, 1975–1987.
- Walpole, A.A., Bowman, J., Murray, D.L. & Wilson, P.J. (2012) Functional connectivity of lynx at their southern range periphery in Ontario, Canada. *Landscape Ecology*, **27**, 761–773.
- Zeller, K.A., McGarigal, K. & Whiteley, A.R. (2012) Estimating landscape resistance to movement: a review. *Landscape Ecology*, **27**, 777–797.

Received 16 August 2016; accepted 10 October 2016

Handling Editor: Timothée Poisot

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Fig. S1. Workflow and system overview of GFLOW.

Appendix S1. Installation and execution procedure for GFLOW and GVIEW.