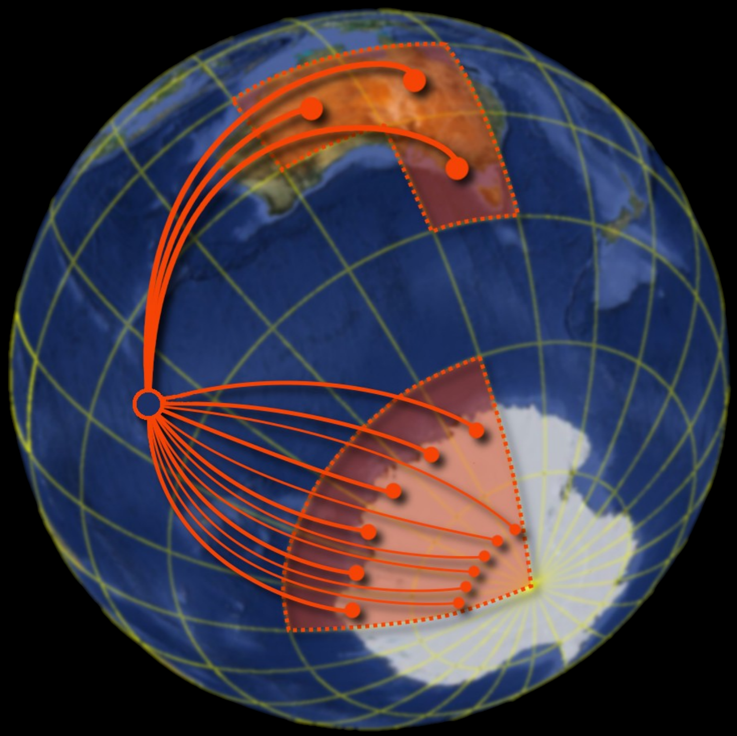
Consistently weighted measures for complex network topologies

Jobst Heitzig, J. F. Donges, Y. Zou, N. Marwan, J. Kurths
Potsdam Institute for Climate Impact Research
Transdisciplinary Concepts and Methods



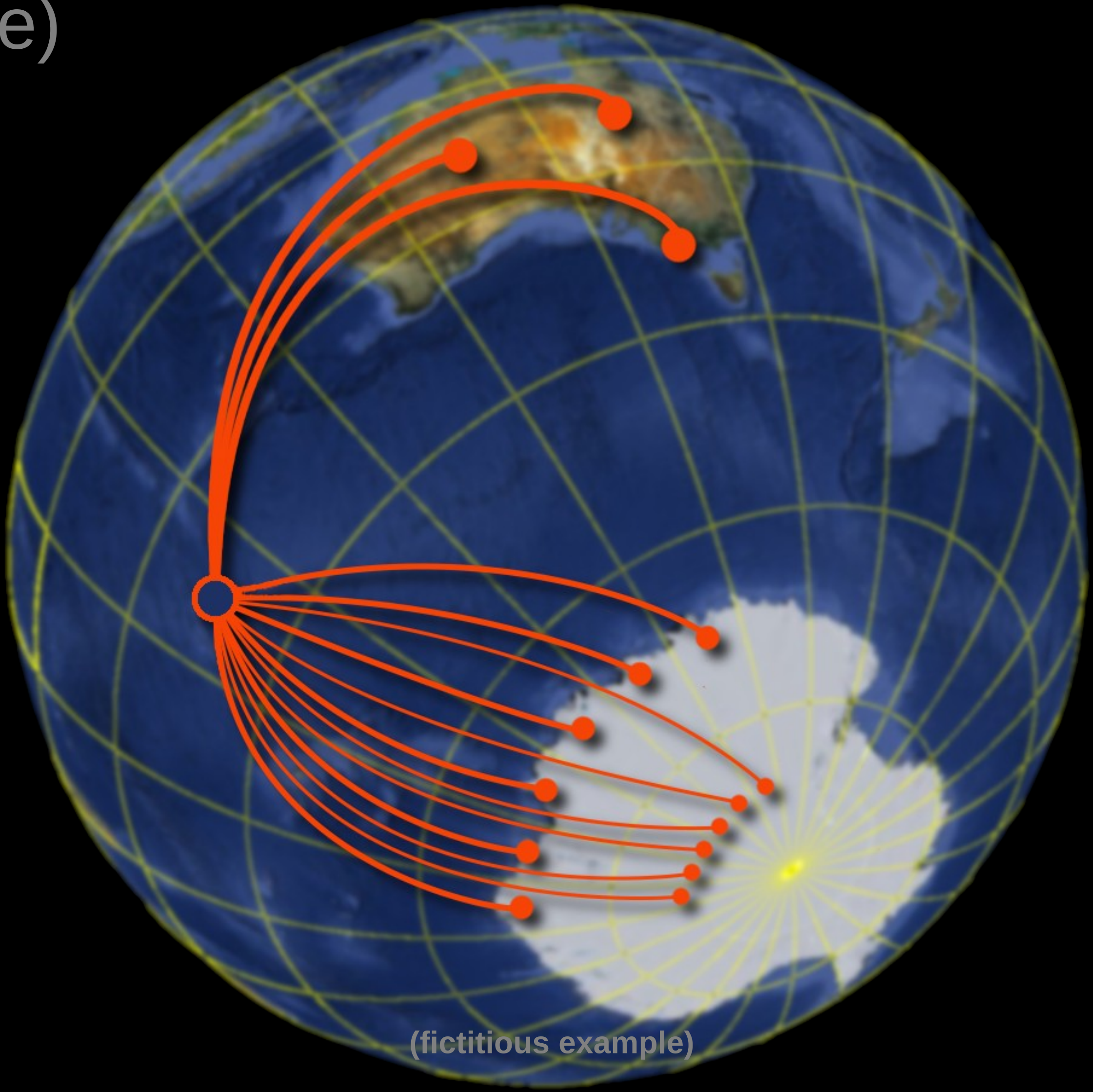
PROGRESS
Potsdam Research Cluster for Georisk Analysis, Environmental Change and Sustainability

POTSDAM INSTITUTE FOR CLIMATE IMPACT RESEARCH
PIK

Motivation: Climate Networks

Nodes represent grid cells,
cell size varies ≈ cos(latitude)

Network measures
are based on counting
(nodes, links, paths...)



Motivation: Climate Networks

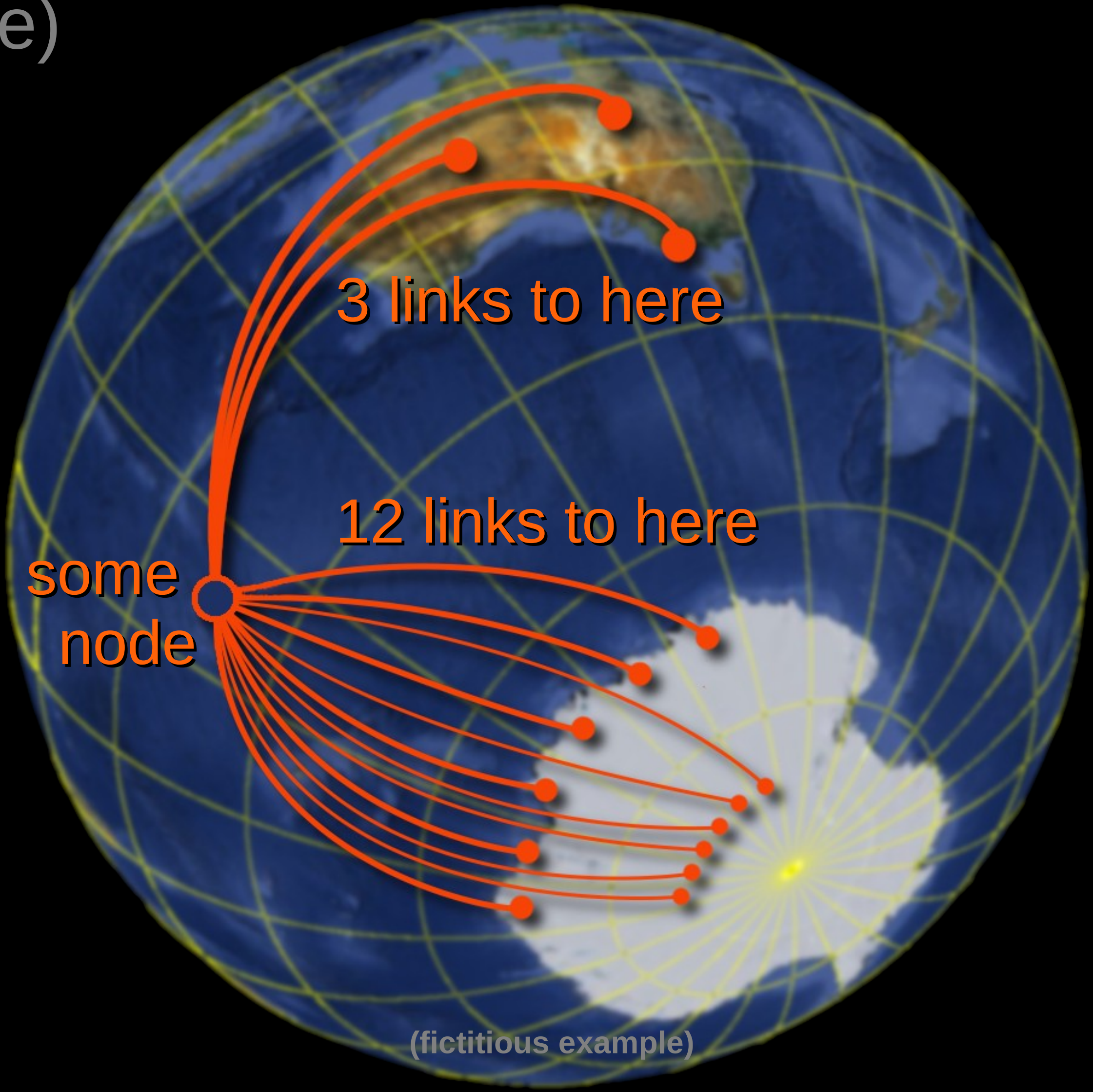
Nodes represent grid cells,
cell size varies ≈ cos(latitude)

Network measures
are based on counting
(nodes, links, paths...)

↓

Polar regions are
over-represented

Results can get biased
or show artificial features



Motivation: Climate Networks

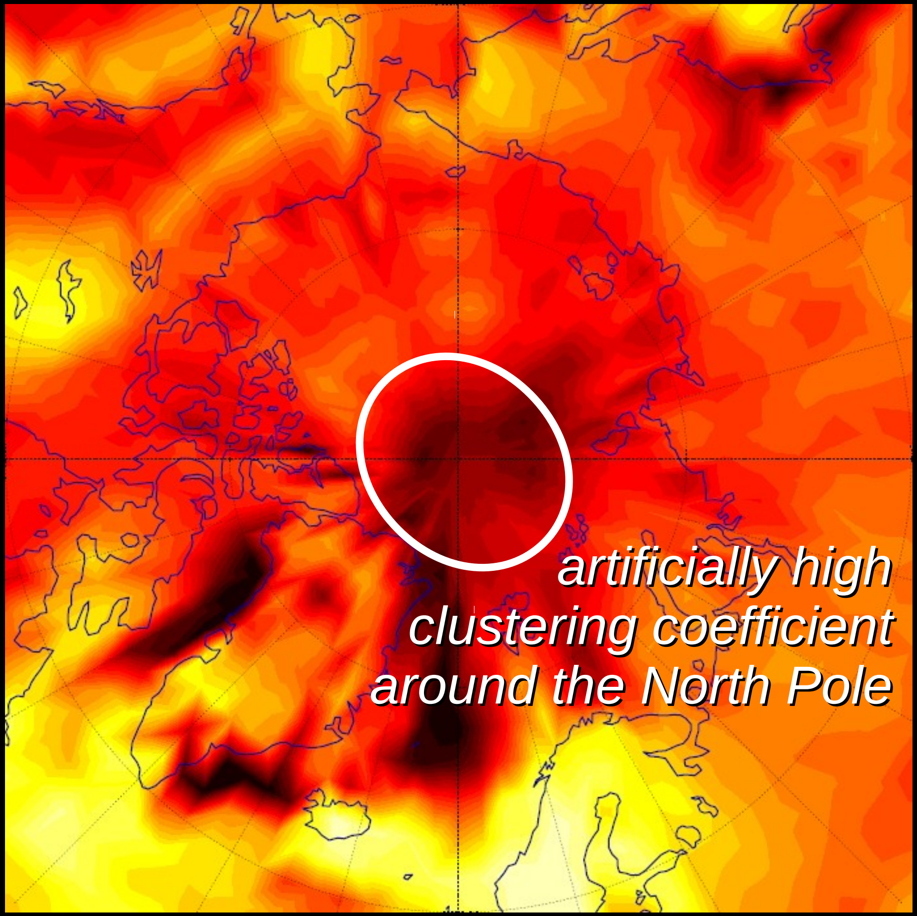
Nodes represent grid cells,
cell size varies ≈ cos(latitude)

Network measures
are based on counting
(nodes, links, paths...)

⬇

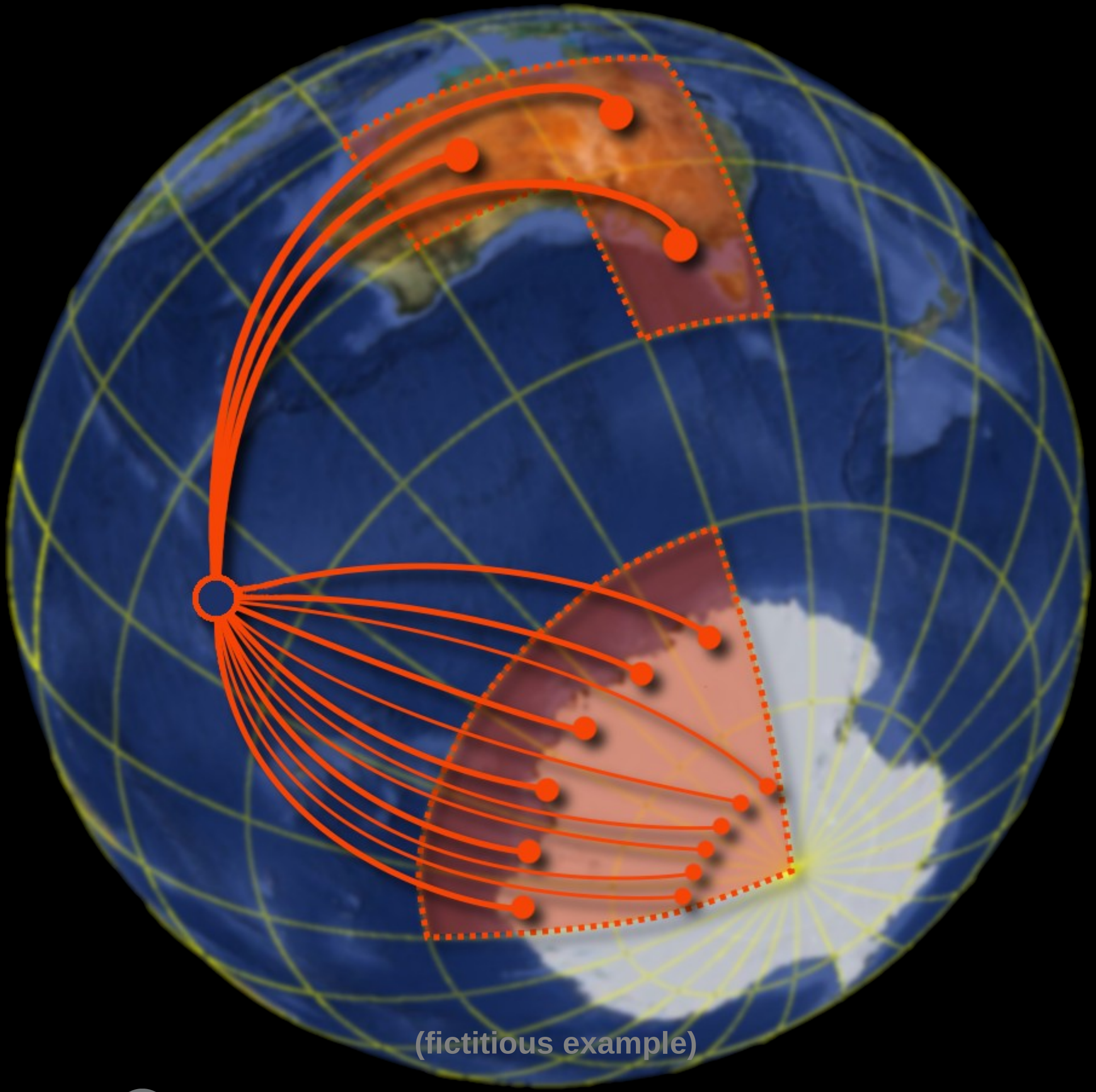
Polar regions are
over-represented

Results can get biased
or show artificial features



Natural idea: Use weights

Cell size → Node weight



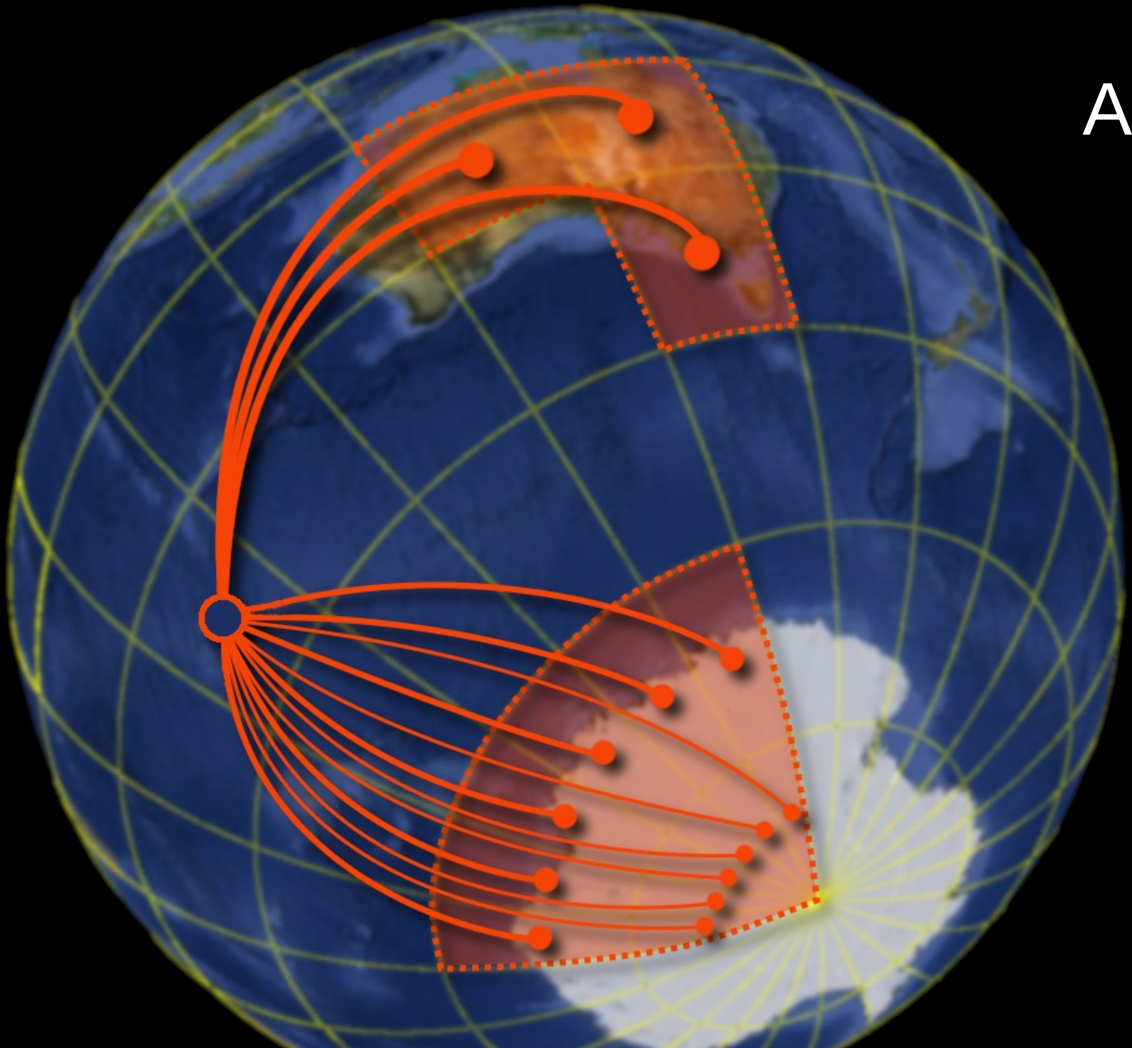
(fictitious example)

Natural idea: Use weights

Cell size ➙ Node weight

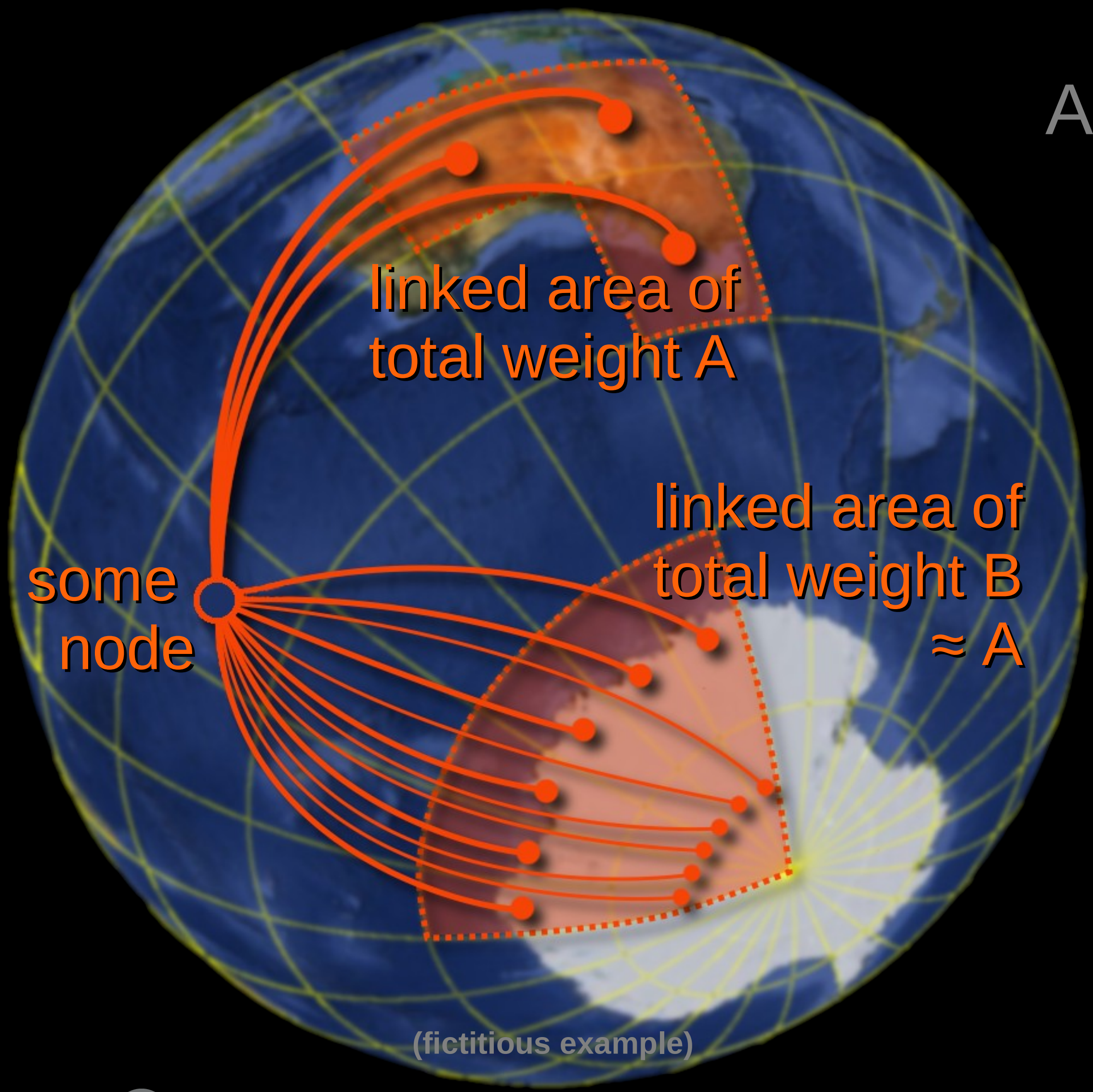
Almost no network measures use *node* weights already

Existing measures using *link* weights don't help



(fictitious example)

Natural idea: Use weights



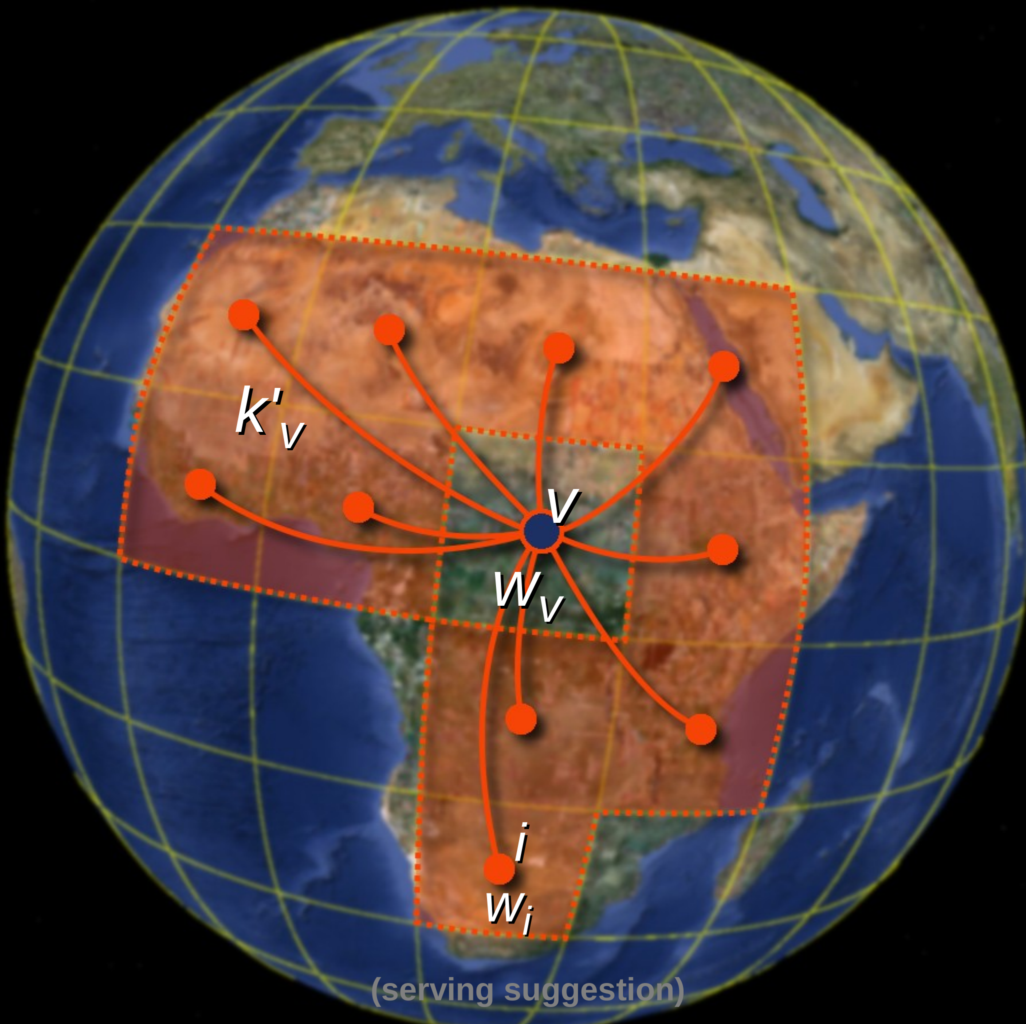
Cell size ➙ Node weight

Almost no network measures use *node* weights already

Existing measures using *link* weights don't help

Find node-weighted versions of measures (degree, clustering coeff., betweenness, spectrum, ...)

Simple example: The "degree" measure

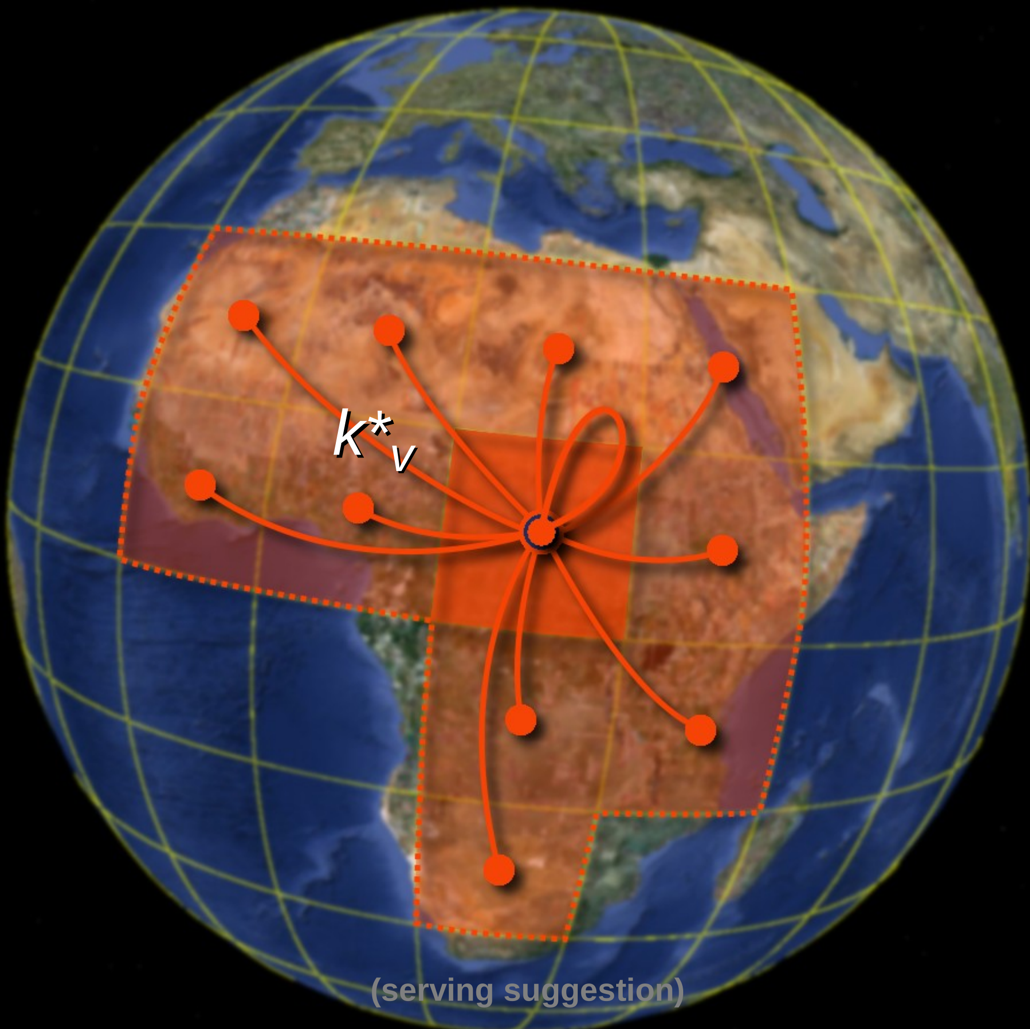


Nodes v, i, …
node weights w_v, w_i, …

Degree:
k_v = no. nodes linked to v

Area-weighted connectivity:
k'_v = sum of w_i
for all i linked to v
(Tsonis et al. 2006)

Simple example: The "degree" measure



Nodes v, i, …
node weights w_v, w_i, …

Degree:
k_v = no. nodes linked to v

Area-weighted connectivity:
k'_v = sum of w_i
for all i linked to v
(Tsonis et al. 2006)

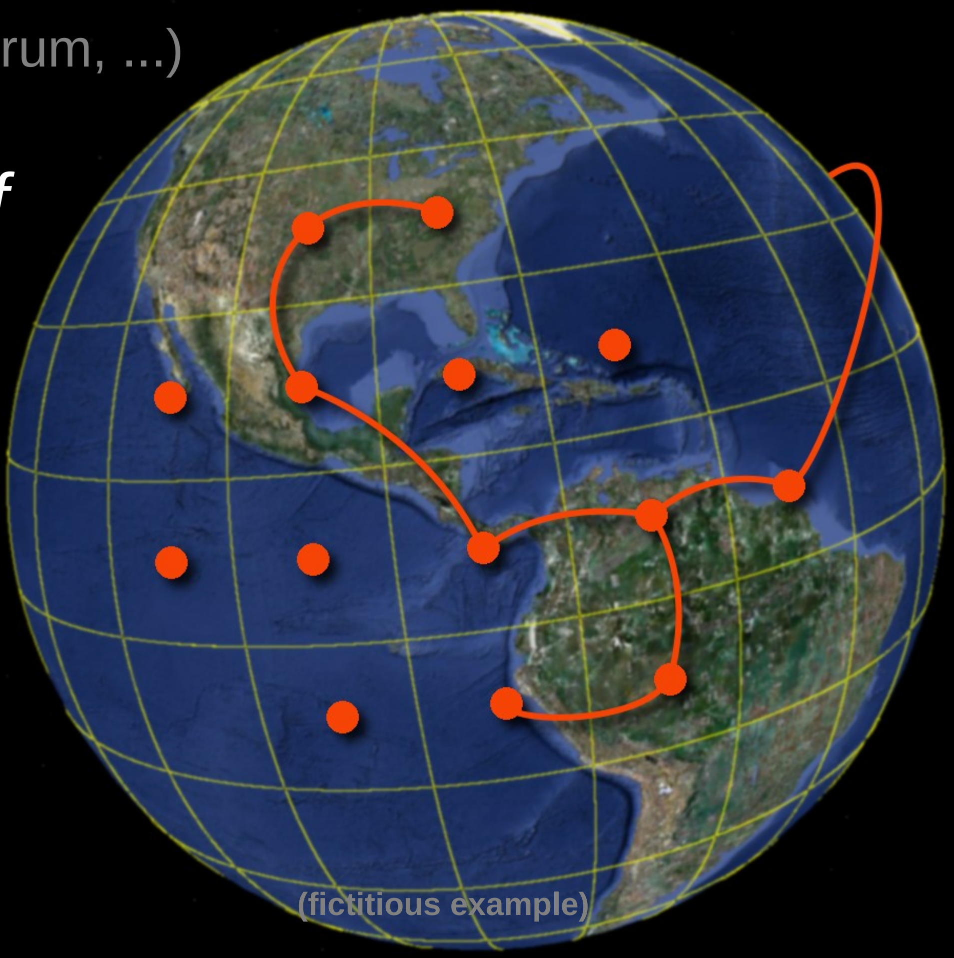
Better version of weighted degree:
$k^*_v = k'_v + w_v$

Why *k** and not *k'*?
And what about more complex measures?

Goal: Find the right way of using the node weights w_i in some given measure f

(degree, clustering coeff., betweenness, spectrum, ...)

Idea: Consider what happens to f when the grid is refined!



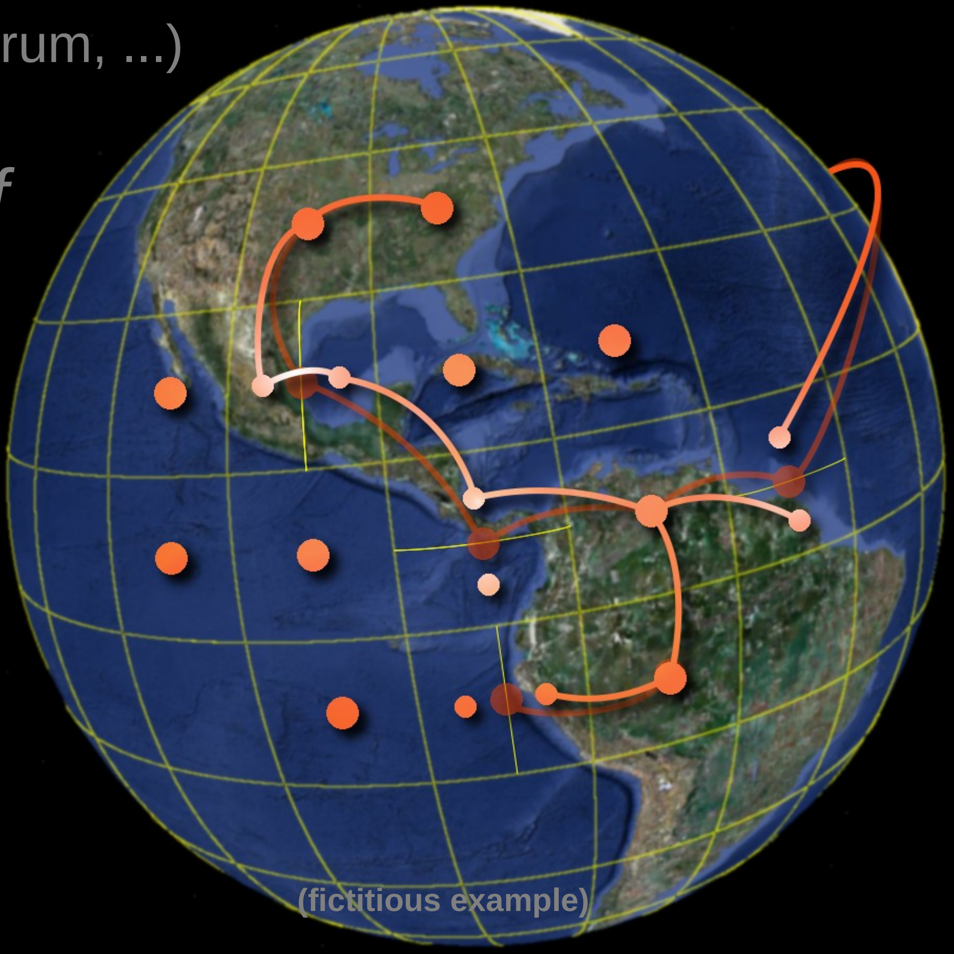
Why *k** and not *k'*?
And what about more complex measures?

Goal: Find the right way of using the node weights w_i in some given measure f

(degree, clustering coeff., betweenness, spectrum, ...)

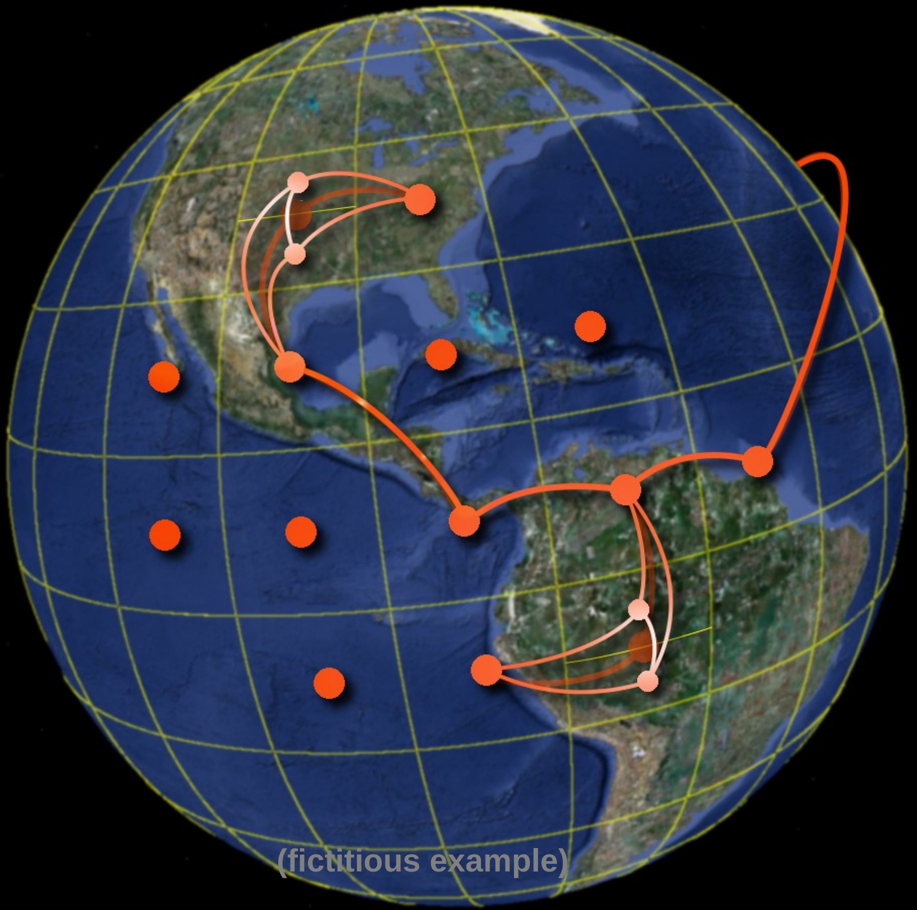
Idea: Consider what happens to f when the grid is refined!

Example:
Under typical refinements,
f should get more realistic →



Redundant refinements / General guideline

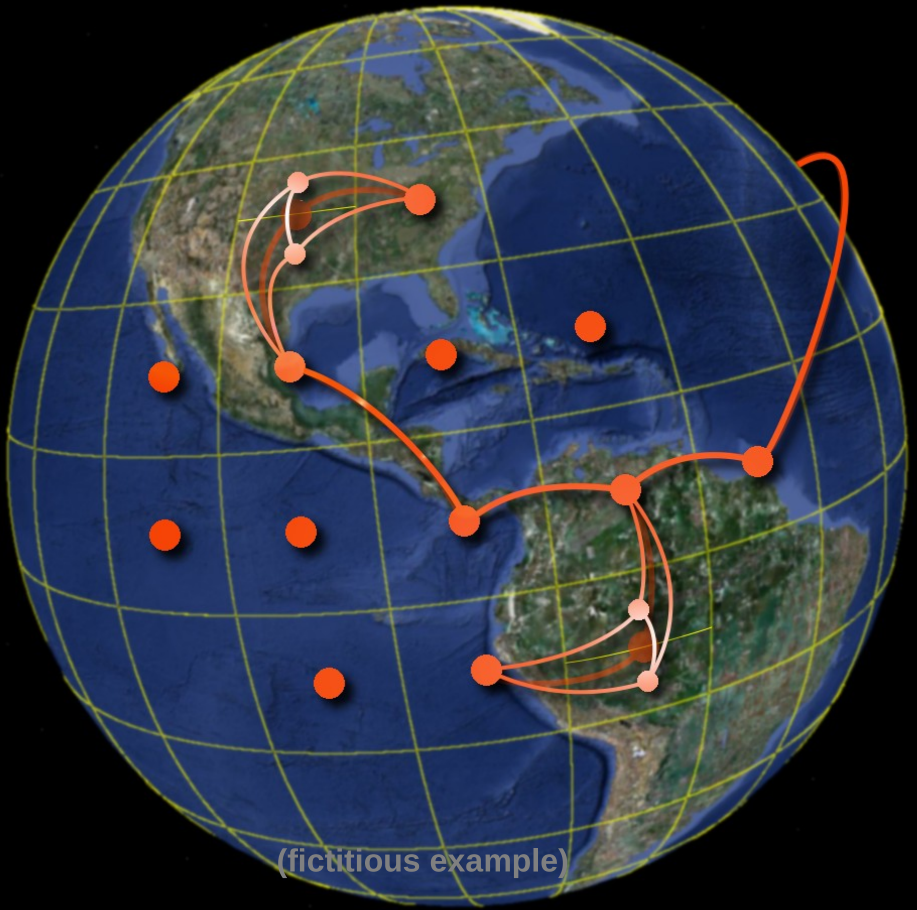
Under “redundant” refinements → *f* should *not* change



Redundant refinements / Guiding notion

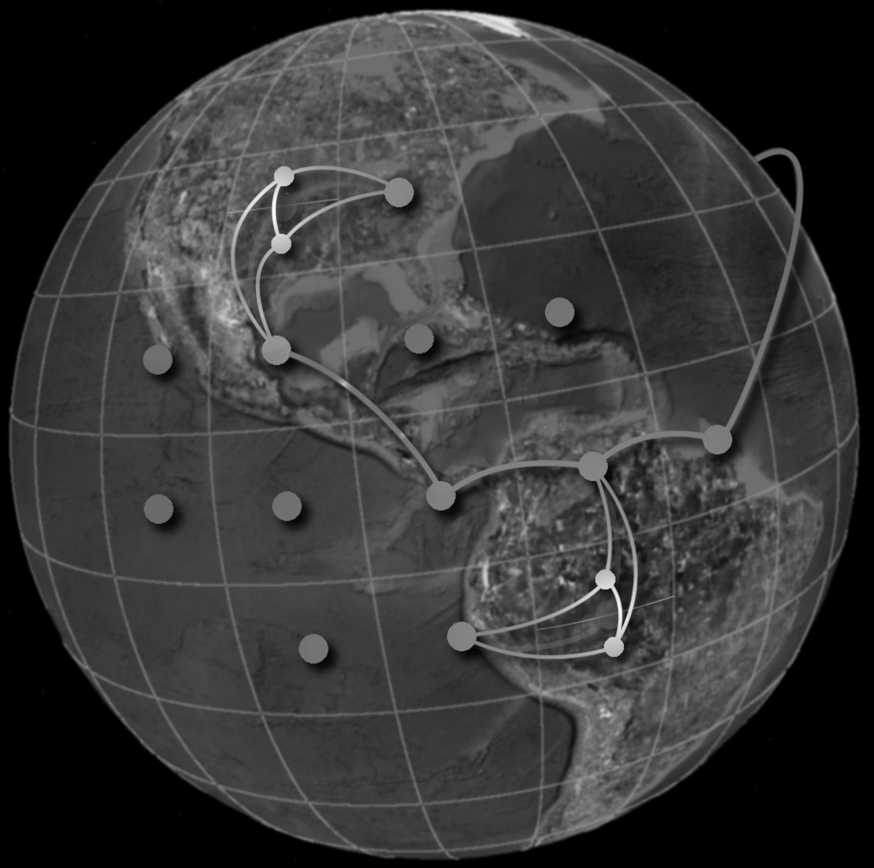
Under “redundant” refinements →
<u>*f* should *not* change</u>

This vague requirement helps to find the weighted formula *f**
for a given measure *f* !



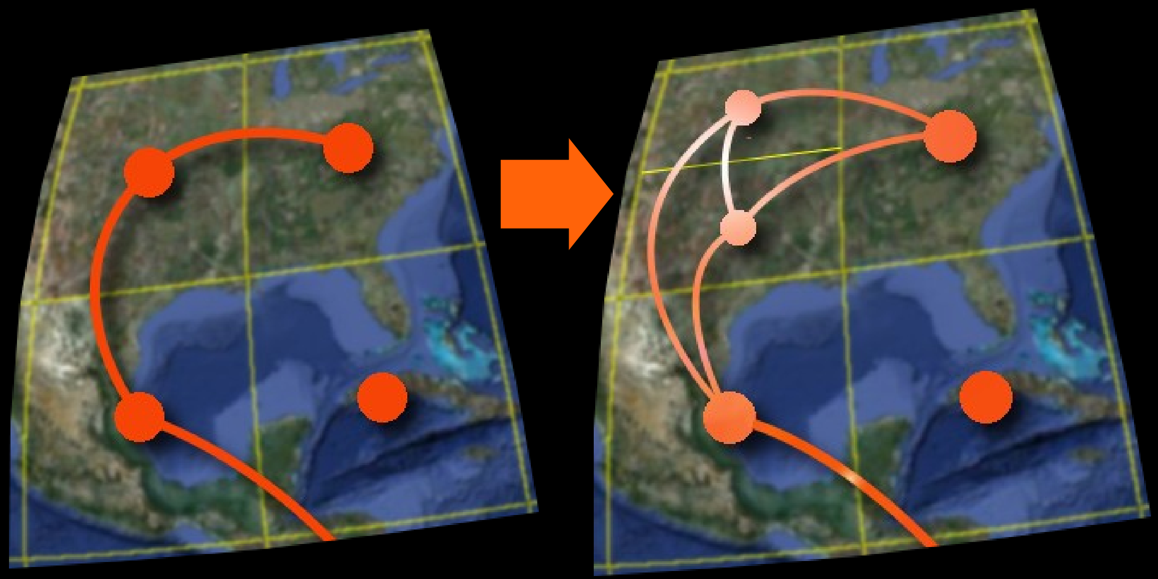
Redundant refinements / Guiding notion

Under *redundant* refinements, →
f should *not* change



This vague requirement helps to find the weighted formula *f** for a given measure *f* !

Guiding notion: Call *f**
“node splitting invariant”
if it doesn't change under this kind of node splitting:



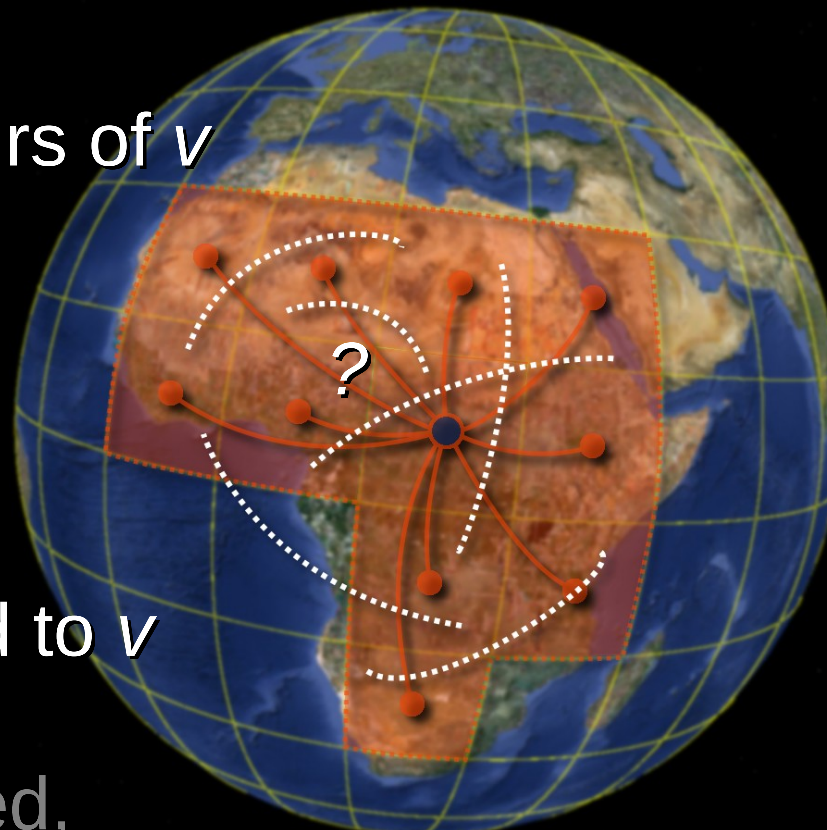
2nd Example: Clustering coefficient

Measures how closely linked the neighbours of v are.

Usual formula:

C_v = rate of links between neighbours of v

$$= \Sigma_i \Sigma_j a_{vi} a_{ij} a_{jv} / k_v (k_v - 1)$$



Node splitting invariant formula:

$$C^*_v = \Sigma_i \Sigma_j a'_{vi} w_i a'_{ij} w_j a'_{jv} / k^*_v k^*_v$$

= link density in the region linked to v

In this, $a_{ij} = 1$ means i and j are linked,
and $a'_{ij} = 1$ means i and j are linked or equal

Useful techniques for formula construction

$$C_v = \Sigma_i \Sigma_j a_{vi} a_{ij} a_{jv} / k_v (k_v - 1)$$

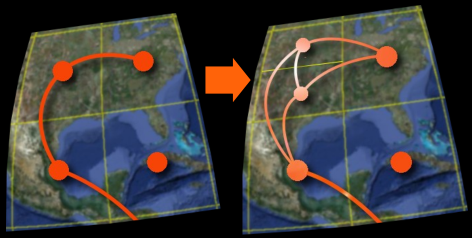
$$C^*_v = \Sigma_i \Sigma_j a'_{vi} w_i a'_{ij} w_j a'_{jv} / k^*_v k^*_v$$

Consider each node a neighbour of itself (e.g. replace a_{ij} with a'_{ij})

Replace edge counts by sums of weight products

Replace node counts by sums of weights

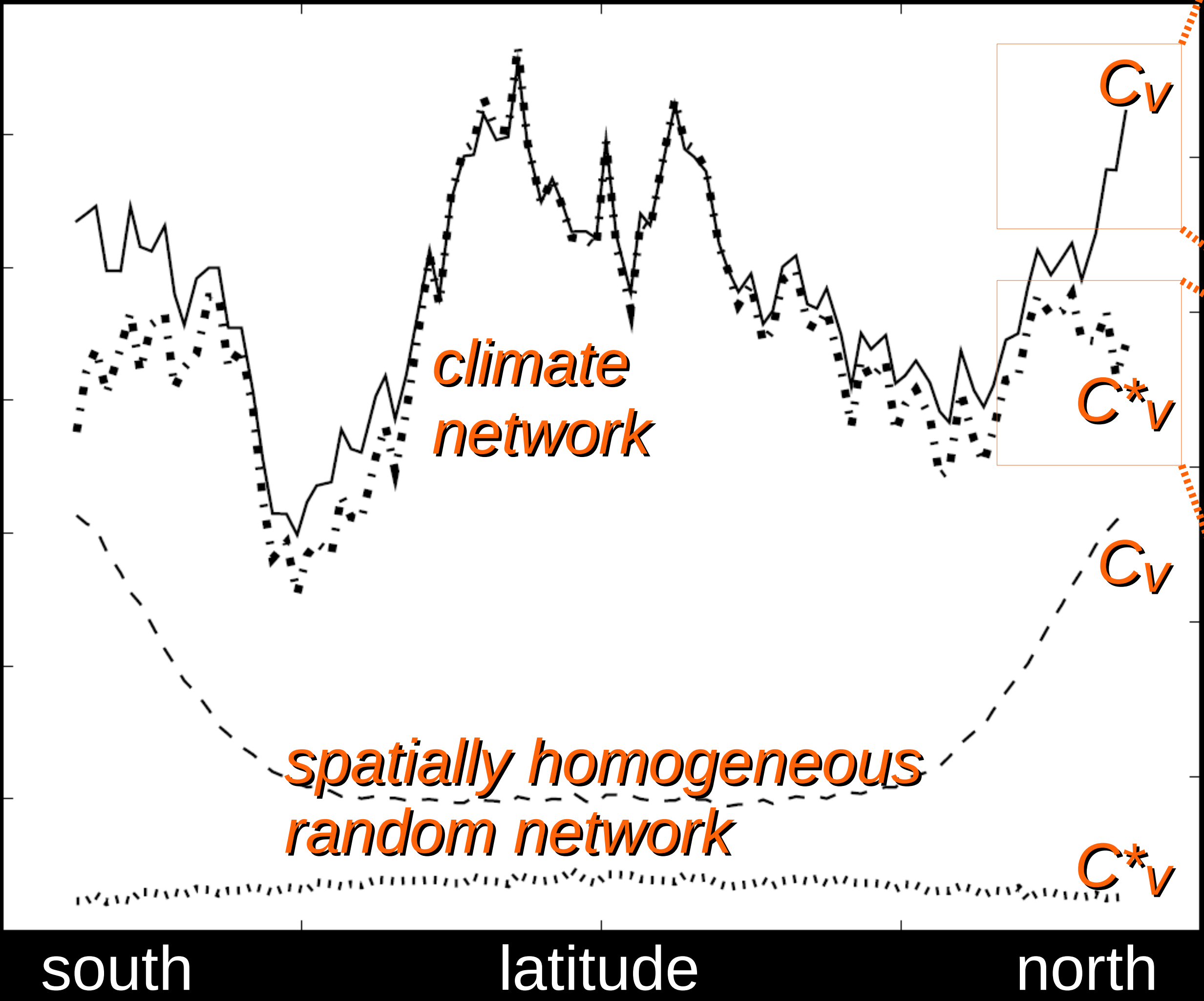
Plug in weighted instead of unweighted measures (k^*_v instead of k_v in this case)

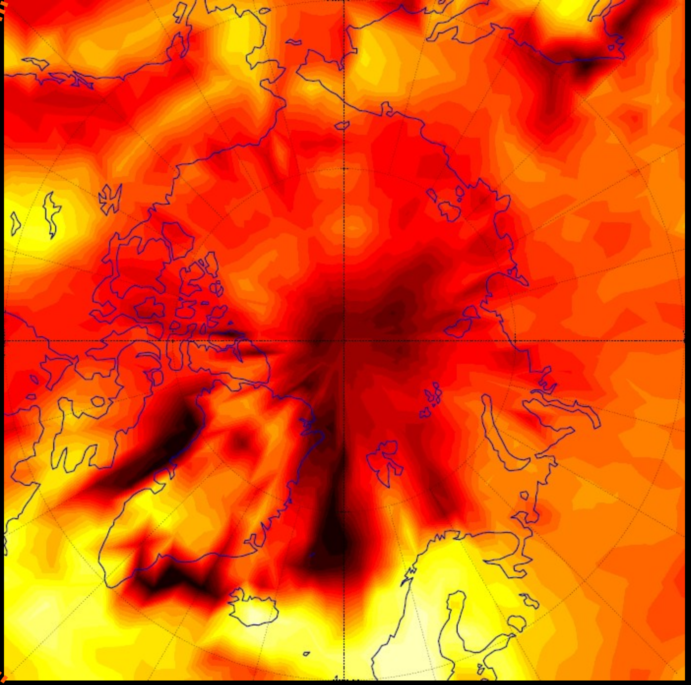


Verify the result is indeed node splitting invariant!

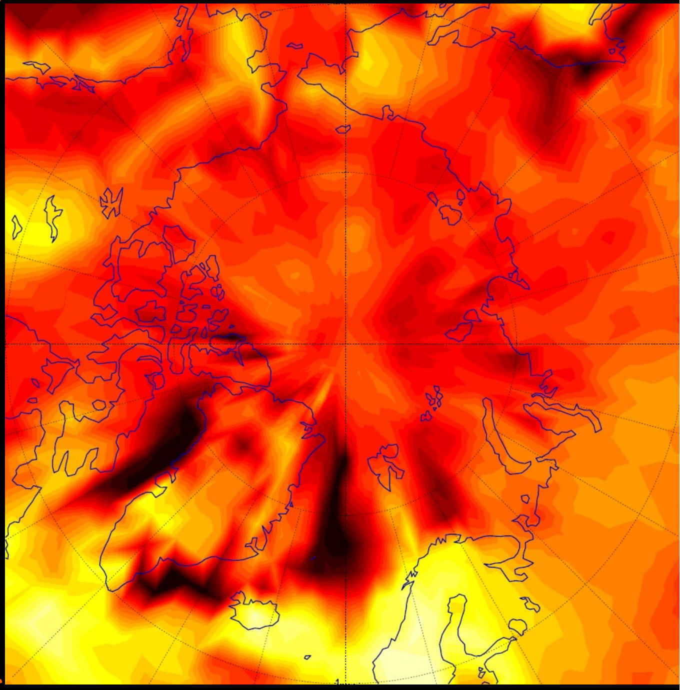
Effect in climate networks

Clustering coefficient averaged by latitude

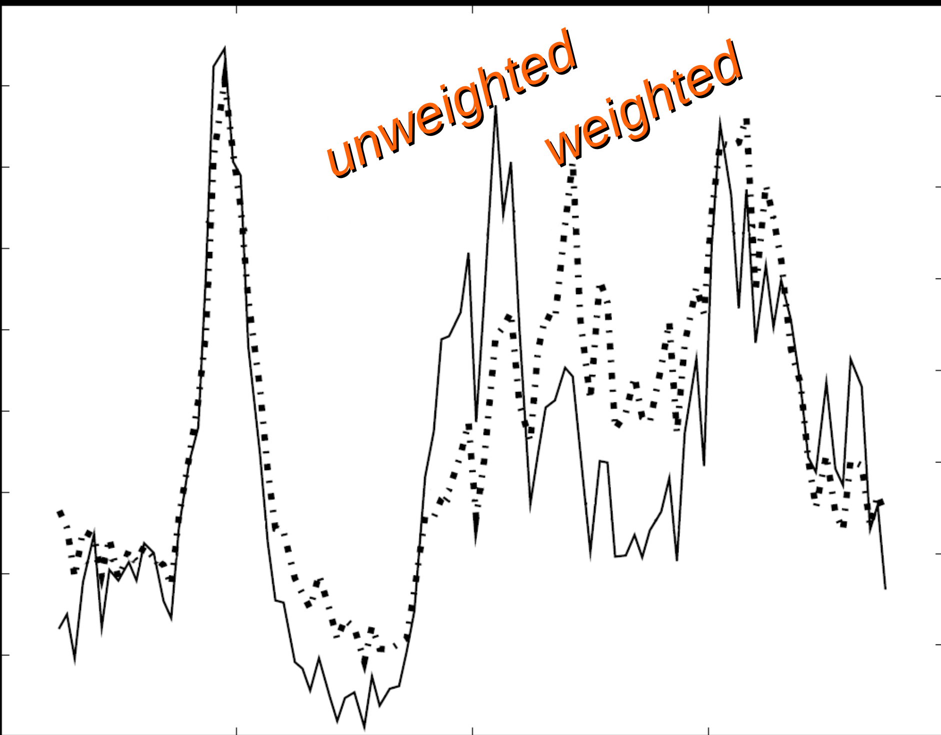




(dark is high)

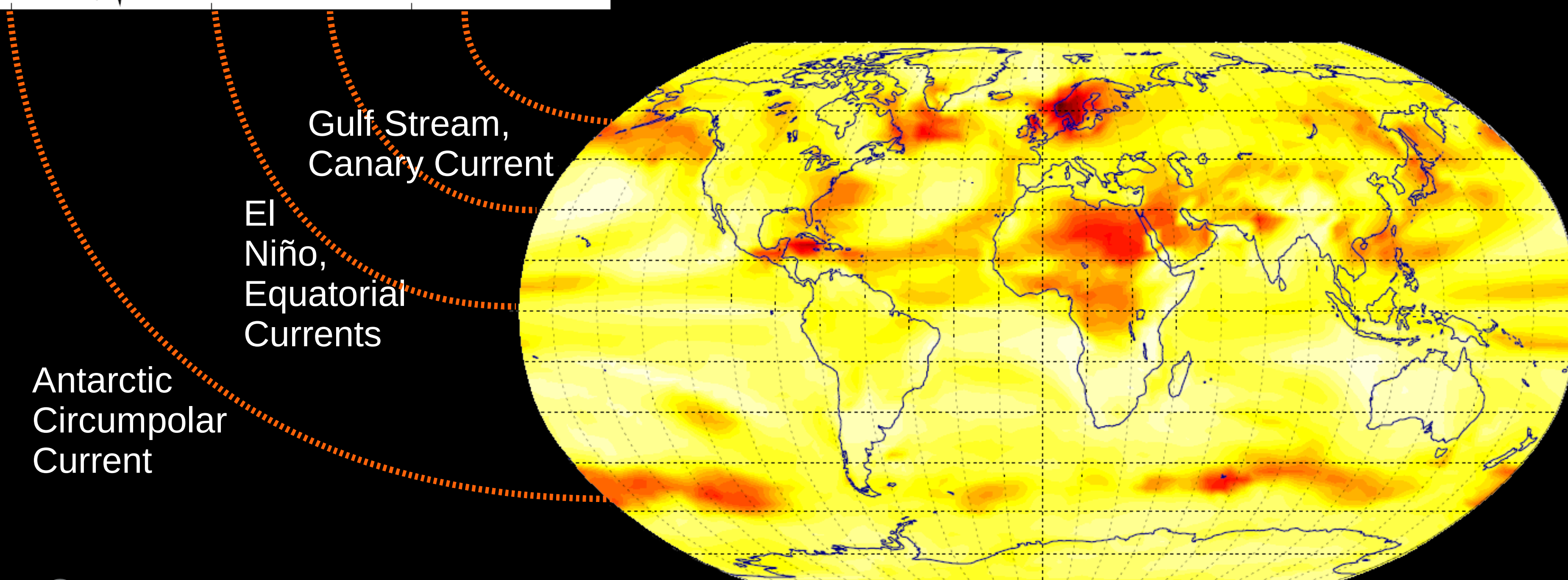


Final example: Newman's random walk betweenness



Measures “importance” of nodes based on Kirchhoff's equations

Unweighted and weighted versions highlight slightly different features



References

J. Heitzig, J. F. Donges, Y. Zou, N. Marwan, J. Kurths (2010), Consistently weighted measures for complex network topologies, under review.

A. A. Tsonis, K. L. Swanson, P. Roebber (2006), Bull. Am. Meteorol. Soc. 87, 585.

Contact

Jobst Heitzig
heitzig@pik-potsdam.de
www.pik-potsdam.de/members/heitzig

PROGRESS
Potsdam Research Cluster for Georisk Analysis, Environmental Change and Sustainability

PIK
Potsdam Institute for Climate Impact Research