The Future scientific challenges for CORDEX: Empirical Statistical Downscaling (ESD)

José M. Gutiérrez, Tereza Cavazos, Jason Evans, Grigory Nikulin, Samuel Somot, Douglas Maraun, Rasmus E. Benestad, Bruce Hewitson, Maria L. Bettolli

Draft version v1 (initial draft from SAT members): 25 November 2021.
v2 version – including comments from POCs until 11 February 2022.
v3 version – final version (21 June 2022)

Preamble

A CORDEX *white paper* describing the scientific challenges in regional climate modelling and setting the basis for the CORDEX science plan and for a better-informed decision-making process at regional scale was made publicly available in May 2021 (Solman et al. 2021). While the first paper focused primarily on dynamical downscaling, here we present a complementary paper focusing on empirical statistical downscaling (ESD) strategies. We describe the ongoing CORDEX ESD work and identify specific challenges, both methodological and more practical, related to the provision of useful information to decision-makers through ESD. A description of the CORDEX general framework (including domain activities and Flagship Pilot Studies, FPS) and links with other initiatives are described in Solman et al. (2021).

Both documents form the basis of the CORDEX Science Plan and have been open for comments so that the community can participate in the development of CORDEX and regional climate science.

Current state and achievements of CORDEX ESD activities

Downscaling methods follow a wide range of approaches (perfect prognosis – PP, model output statistics – MOS, hybrid PP-MOS and bias adjustment, weather generators, etc.) which are applied individually or in combination to produce regional/local information from global or regional model outputs. Most of the work developed in the framework of CORDEX can be broadly classified in the following groups of activities:

Perfect prognosis (PP) – relating large-scale GCM predictors to local response. PP methods have been developed and applied in different CORDEX domains (including Europe - EUR, South America - SAM, East Asia - EAS, and South Asia - WAS) and FPSs, typically at a regional (national or subnational) scale, considering station data and focusing on a few variables (e.g. precipitation and temperature). Some intercomparison studies have been conducted assessing the performance of the methods with perfect (reanalysis) predictors (e.g. over EUR: Maraun et al., 2015; Vaittinada et al., 2015; Gutiérrez et al., 2019; Maraun et al., 2019; and SAM: Bettolli et al., 2021). Recent developments using machine learning techniques (in particular deep learning) have allowed continental-scale applications using gridded observations more comparable with the RCM experimental framework (Vandal et al., 2019; Baño-Medina et al., 2021). The typical application of these methods is to use the relationships as found in observations to predict/simulate an expected/stochastic value from a GCM predictor, or to infer a local response conditioned in some way by the GCM. In addition, downscaling can be approached by matching the large-scale conditions with local conditions for each time-step ('downscaling weather') or downscaling the parameters of probability density functions (pdfs) directly ('downscaling climate'; Benestad, 2021).

The selection and transformation of predictors have a major influence in the configuration of ESDs (Cavazos and Hewitson, 2005; Gutiérrez et al., 2019) and a variety of approaches have been used, including pre-processing approaches such as the use of EOF and neural networks to capture main spatial modes of variability (features). Some machine-learning approaches build on spatial features learn automatically from data via, e.g. convolutional neural networks (Baño-Medina et al., 2021), self-organizing maps (SOMs) (Cavazos 2000; Skific et al., 2010).

Hybrid PP-MOS for predictor data framing. The Hybrid PP-Model output statistics (MOS) approach involves both reanalyses and GCM simulations in the calibration of the downscaling methods, and under the "common EOF" approach, both GCM and reanalysis information are used to calculate the EOFs thus bringing extra information in the calibration phase. Whereas MOS accounts for systematic biases in the simulated output, common EOFs ensure that the exact same covariance structure used in the calibration of the downscaling methods is utilised when downscaling the GCM output. They also reveal biases in the simulations and can facilitate a bias correction of the GCM results (Benestad et al., 2001).

MOS Bias adjustment – adjusting GCM/RCM model outputs to local scale. The development of bias adjustment methods for GCM/RCM model outputs according to regional/local observations is an active field of research boosted by the increasing demand and use of these products by the impact and adaptation communities. RCM bias adjustment can be considered a hybrid method in the sense that it combines both dynamical and statistical downscaling. A variety of applications and intercomparison studies have been undertaken in many CORDEX domains (see the CORDEX publications list[1]) and actionable datasets of adjusted CORDEX RCM ensembles have been produced and made publicly available (e.g. McGinnis and Mearns, 2021). The CORDEX archive design (and the data reference syntax, DRS) have been expanded to archive CORDEX bias adjusted[2] results in the Earth System Grid Federation (ESGF) and some adjusted simulations are already available for several domains[3]. A discussion on bias adjustment including assumptions and limitations can be found in Maraun et al. (2017) and the recent IPCC report (Doblas-Reyes et al., 2021).

[1] https://cordex.org/publications/peer-reviewed-publications

[2] http://is-enes-data.github.io/CORDEX_adjust_drs.pdf

[3] https://cordex.org/data-access/bias-adjusted-rcm-data

RCM emulators – simulating RCM outputs using statistical methods. Emulators are a more recent hybrid approach where computationally-cheap statistical methods (including advanced machine learning such as neural networks) are used to learn the RCM downscaling function in order to be able to mimic high-resolution RCM outputs using coarser driving GCM predictors. The final aim of RCM emulators is to emulate at low cost partially or in full the spatio-temporal complexity of the original high-resolution RCM for chosen variables of interest (e.g. to emulate convection-permitting nested to standard CORDEX simulations). With respect to standard ESD, this hybrid downscaling approach allows to work over areas where high-quality observations are not available and to take into account climate-dependent evolution of the large-scale/small-scale relationship as learning in future simulations is possible. The approach furthermore allows to include feedbacks and nonlinear local responses to large-scale climatic changes as represented by RCMs or convection-permitting regional climate models (CPRCMs) which may not be easily derived from pure ESD approaches. Pioneer work includes the emulation of daily surface temperature maps at 12 km over Europe (Doury et al., 2022). Besides, ta first experimental protocol for emulator intercomparison has been defined in the framework of CORDEX FPS on convection[4]. So far, within CORDEX, there is no standardization of the naming for the RCM emulators or the files produced by emulators. This work will be needed in the coming years.

Future Challenges

1. Advances in ESD methodological aspects

There are a number of methodological challenges for the different ESD approaches including the following.

Generalization of the ESD methods to future climates. One of the main requirements for downscaling methods in the context of climate change is that the relationships established in current climate conditions generalize to future climates. Problems to be avoided are overfitting (in particular machine learning-based methods with a large number of parameters), or using predictors not carrying information on the climate change signal. Beside specific studies carried out in the framework of CORDEX to assess this problem for different approaches, there are some ongoing intercomparison experiments[5] to test generalization in ensembles of ESD methods covering different approaches and techniques. Future coordinated activities will be required to gain comprehensive understanding on this challenge. This problem is particularly relevant in the case of bias adjustment methods and further research is needed to understand the effects on trends.

Multivariate Aspects. Standard ESD approaches, e.g., using PCA, represent the local covarying behaviour imposed by large-scale predictors. However, residual spatial, temporal and multivariable dependence have to be explicitly modelled by a multivariate noise model (e.g. Maraun et al., 2015; Maraun and Widmann, 2018).

---

[4] https://docs.google.com/document/d/1266C1tUgrXV-cwxvHy8mY1XRVSsPnXUBV9nUsf2q940

[5] http://www.value-cost.eu/validation/#Experiment_2a

Weather generators are designed to provide such models, and if driven by large-scale predictors these can serve as sophisticated PP downscaling methods (e.g., Maraun and Widmann, 2018; Doblas-Reyes et al., 2021). The underlying models comprise, for instance, simple Richardson-type weather generators, truncated Gaussian models, generalised linear models, spatial Poisson cluster models, random cascade models, and non-homogeneous hidden Markov models. Given the complex dependence between different locations and variables, these models may assume a complexity that is computationally challenging. A fundamentally limiting factor is the lack of a closed analytical multivariate distribution beyond the multivariate normal distribution. In particular, the dependence of extreme events (in different variables or at different locations) is methodologically challenging. Here, pair copula constructions, also called vine copulas, have shown potential for decomposing multivariate random variables (e.g., Aas et al., 2009). There are some approaches that may address some of these concerns, such as using the analog method or PCA to represent a group of stations in the predictands and downscaling the parameters of pdfs or aggregated statistics (Benestad, 2021). In the case of Bias Adjustment, different resampling techniques have been proposed to impose multivariate consistency (e.g. Vrac 2018).

2. Advances in intercomparison/validation frameworks

If ESD is to be used to inform decision-makers and for climate change adaptation, it is necessary to evaluate how the strategies and methods perform when applied to ensembles of GCMs. In addition to the traditional cross-validation, there is a need to assess their ability to reproduce historical trends and interannual variability, and check the stationarity assumption. The intercomparison of different approaches and techniques is fundamental for understanding and communicating the advantages and limitations of the different methods. There have been a few intercomparison experiences so far considering perfect predictors (Maraun et al., 2015; Vaittinada et al., 2016).

The VALUE initiative (Maraun 2015) developed a comprehensive intercomparison framework and produced results for over fifty ESD methods in perfect predictor conditions over Europe[6]. Some ongoing activities are expanding the intercomparison over Europe and South America using GCM projections to analyze the uncertainty of the downscaled results[7]. Building on these experiences, an intercomparison experimental protocol is being developed for CMIP6[8]. Other challenges are the use of weighted (constrained ensembles) and unweighted mean ensembles, and regional versus gridpoint downscaling. These topics were analyzed by Colorado-Ruiz et al., (2018) in a study in CORDEX-CAM (Central America, the Caribbean, and Mexico) where they evaluated two versions of the reliability ensemble averaging (REA; Giorgi and Mearns, 2002; Xu et al., 2010) technique at different spatial scales for weighted and unweighted GCM ensembles from CMIP5. Other approaches, such as pseudo-global warming or pseudo-observation use model (GCM or RCM) outputs for both predictors

---

[6] See the IJOC special issue, https://rmets.onlinelibrary.wiley.com/toc/10970088/2019/39/9

[7] http://www.value-cost.eu/validation/#Experiment_2a

[8] https://docs.google.com/document/d/1det0PnLcAILQdOta7AJopADz0k3_QgX1hJ06OrhR1-w

and predicands to develop either dynamical or statistical models to analyze scientific problems (e.g. generalization capability of the dynamical or statistical methods to extrapolate results in climate change conditions; Erlandsen et al., 2020).

Questions

- How do we ensure/evaluate that the GCM predictors simulate credible future predictors?
- How do we ensure/evaluate that the statistical model includes all relevant predictors to represent long-term changes in a realistic way?
- How do we ensure/evaluate that the statistical model is credible under extrapolation to unobserved regions of the predictor space?
- How do we take into account observational scarcity/uncertainty in the evaluation process and in the design of coordinated ESD experiments?
- Can we produce suitable evaluation approaches targeted to specific aspects of bias adjustment (downscaling, temporal structure, etc.) ?
- Can we produce comprehensive intercomparisons to understand and communicate the benefits and limitations of ESD approaches and techniques?
- How do we measure the added value of downscaling in evaluation experiments?, what are the appropriate metrics taking into account the diversity of results/outputs?

3. Machine Learning for ESD

Machine learning techniques have been used and intercompared in downscaling applications since early work (von Storch and Zorita, 1999). Recently, this field has gained renewed attention boosted by major breakthroughs obtained with deep learning (DL) models (Reichstein et al. 2019). The advantage of DL resides in its ability to extract high-level feature representations in a hierarchical way due to its (deep) layered structure. In particular, convolutional neural networks (CNNs) have gained great attention in spatiotemporal problems due to their ability to learn spatial features from data. There have been some attempts to test the application of these techniques for ESD, including simple illustrative examples of super-resolution approaches to recover high-resolution (e.g. precipitation) fields from low-resolution counterparts with promising results, spatially consistent weather generators, perfect prognosis applications using different methods and architectures (Baño-Medina, 2021) and RCM emulators (Doury et al., 2022). These complex (in many cases off-the-shelf) models are usually seen as black boxes, generating distrust among the climate community, particularly when it comes to climate change problems. Recently, Reichstein et al. (2019) outlined this problem and encouraged research towards the understanding of deep neural networks in climate science.

Questions

- How do we ensure/evaluate that the automatization of predictor selection and feature extraction captures the right physical phenomena needed for downscaling?

- How do we ensure/evaluate that machine learning methods produce plausible projections generalizing to future climates?
- Can we advance in the understanding of machine learning methods to gain interpretability of results?
- Can we build RCM emulators suitable for certain tasks (e.g. filling-up temporal gaps in very-high-resolution CPRCM runs, filling-up the SCEN/GCM/RCM matrix, to better explore the natural variability related uncertainty, create large ensembles over areas poorly covered by RCM runs)?

4. Distillation of actionable information

A variety of climate projection sources are available providing information for regions including ensembles of GCMs – including recent high-resolution simulations – further downscaled by dynamical and statistical downscaling thus generating different layers of information relevant for regions. Despite the increasing availability of these ensembles, the provision of actionable and defensible information about regional climate change is yet to be operationalized and involves many aspects on what information means, the role of context, the region to be analyzed, etc. (a comprehensive discussion can be found in the recent IPCC report; Doblas-Reyes et al., 2021).
Assessing the added value of the different approaches and techniques is of key importance to understand sensible distillation approaches for different uses and requirements. This is a challenge for the CORDEX community which would require close collaboration with all users of regional information. There are already ongoing co-production experiences in the context of CORDEX (e.g. estimating intensity-duration-frequency curves from daily data; Benestad 2021) but there has been a lack of coordination to develop and promote good practices. Also, besides the global model outputs, actionable information needs to include both dynamical and empirical-statistical downscaling since they have different strengths and weaknesses and make use of information derived from different sources.

The distillation challenge is one of the many aspects involved in the provision of regional information for society, so CORDEX will contribute to relevant initiatives focusing on a broader perspective, such as the Regional Information for Society (RIfS) WCRP core project.

Questions

- What is the influence of large-scale errors for the specific downscaling context (garbage in/garbage out)?
- What is the influence of present-day biases in relevant processes and trends on the climate change signal (emergent constraints)?
- How to reduce the complexity of climate change information (multi-model ensembles) by informed sub-sampling?
- What climate information is used by stakeholders and decision-makers, how do they use it, and is its use consistent with the best science?
- How can we estimate the added value of the different approaches providing regional information, and what are the appropriate metrics?

- To what extent does multi-statistical downscaling model ensembles make sense (they may have different inherent limitations)?
- How to understand/explain the differences or even contradictions between results from different ESD methods, and between ESD and RCMs and between ESD and GCMs?
- How to identify and quantify vulnerability to sationarity?

5. Data and Infrastructure

One of the major successes of CORDEX has been its contribution in making available a set of coordinated dynamical downscaling simulations covering almost all land areas of the world, built mainly on the Earth System Grid Federation (ESGF) infrastructure. Standard CORDEX publishing protocols (e.g., format, variables, time periods and archival conventions) were established as part of the experimental framework and they have been widely used for storing and making available the RCM simulations across the different domains, thus facilitating data access to the vulnerability, impacts and Adaptation (VIACS) community. However, a similar protocol is yet to be completed for ESD. Some difficulties are the specific attributes for ESD methods and the lack of common simulation domains (most ESD results are in-house applications over specific national or subnational domains). ESD datasets would also require comprehensive metadata to describe the methodological approach and the underlying assumptions. Understanding the data and infrastructure requirements to support robust ESD datasets through CORDEX coordinated experiment design is a major challenge for the future.

Questions

- How to extend the CORDEX DRS and archive specification to include ESD results?
- Is ESGF a suitable home for ESD coordinated activities (e.g. national or subnational domains)?
- How to best define and deploy comprehensive ESD metadata?
- How to store downscaled results for large multi-model ensembles gridded over a selected region that makes it practically possible to distill information in a fast and flexible way?

Suggestions

The extension of the CORDEX Data Reference Syntax (DRS)[9] and archive specification for ESD methods needs to be discussed further in the framework of CORDEX and other initiatives; a preliminary (but incomplete) protocol is already available to serve as the basis for this work. This protocol should also be suitable for bias adjustment approaches and align with ongoing efforts in CORDEX-Adjust[10] and other initiatives, such as ISI-MIP which already stores and distributes ensembles of bias adjusted data. The regular ESGF grid domains (e.g. EUR-11i) could be used to archive those ESD simulations which cover a large portion of the domain. Specific domains could be used

[9] https://cordex.org/experiment-guidelines/experiment-protocol-rcms/

[10] https://cordex.org/data-access/bias-adjusted-rcm-data

for special regions (or point-based) simulations (although minimum requirements should be defined for the suitability of archiving at ESGF). In terms of efficient data storage for statistical downscaling purposes, storage of model data at their effective resolution or storage of EOFs could be used to compress information also facilitating "climate downscaling" approaches.

Last mile on bridging climate science with society needs

The need to establish methods and tools to deliver information for climate change adaptation is the main concern of other initiatives, in particular the WCRP RifS. CORDEX should be coordinated with them in order to be the foundation for developing the understanding of the downscaling added value and explaining the differences and contradictions within downscaling approaches and between downscaling and other approaches, so that a larger distillation community (which is necessarily in partnership with users and user contexts) can have confidence in the information derived from the CORDEX work in the broader activities of distillation.

The challenge here is to establish an open collaboration framework with other related WCRP initiatives with a clear plan of who will do what, so future developments are efficient and coordinated. A central concern is to use the best available information in the right way, and communicate an understanding of what the numbers presented to decision-makers really represent. One example to the contrary is to rely on just one RCM realization.

Most of these scientific challenges are broad and require inter and transdisciplinary collaboration of regional climate modelers with global climate modelers, statisticians (Benestad et al., 2017), communication experts, stakeholders and local experts.

References:

Baño-Medina J, Manzanas R and Gutiérrez JM (2021) On the suitability of deep convolutional neural networks for continental-wide downscaling of climate change projections. Clim Dyn 57, 2941–2951 (2021).
https://doi.org/10.1007/s00382-021-05847-0

Benestad RE (2001) A comparison between two empirical downscaling strategies Rasmus E. Benestad. International Journal of Climatology, 21, 1645-1668

Benestad RE et al. (2021) Testing a simple formula for calculating approximate intensity-duration-frequency curves. Environ. Res. Lett. 16 044009.
https://doi.org/10.1088/1748-9326/abd4ab

Benestad R, Sillmann J, Thorarinsdottir T et al. (2017) New vigour involving statisticians to overcome ensemble fatigue. Nature Clim Change 7, 697–703.
https://doi.org/10.1038/nclimate3393

Benestad RE (2021) A Norwegian Approach to Downscaling, Geosci. Model Dev. Discuss. [preprint], https://doi.org/10.5194/gmd-2021-176

Bettolli ML, Solman SA, da Rocha RP et al. (2021) The CORDEX Flagship Pilot Study in southeastern South America: a comparative study of statistical and dynamical downscaling models in simulating daily extreme precipitation events. Clim Dyn 56, 1589–1608. https://doi.org/10.1007/s00382-020-05549-z

Cavazos T (2000) Using self-organizing maps to investigate extreme events: An application to wintertime precipitation in the Balkans. J. Climate, 13, 1718-1732

Cavazos T**,** Hewitson BC (2005) Performance of NCEP variables in statistical downscaling of daily precipitation. Climate Research, 28, 95-107. http://doi:10.3354/cr028095

Colorado-Ruiz G, Cavazos T, Salinas JA, De Grau P, Ayala R (2018) Climate change projections from Coupled Model Intercomparison Project phase 5 multi-model weighted ensembles for Mexico, the North American monsoon, and the mid-summer drought region. Int. J. Climatol. 38(15), 5699-5716, https://doi.org/10.1002/joc.5773

Erlandsen HB, Parding KM, Benestad R, Mezghani A, Pontoppidan M (2020). A Hybrid Downscaling Approach for Future Temperature and Precipitation Change, Journal of Applied Meteorology and Climatology, 59(11), 1793-1807.
https://journals.ametsoc.org/view/journals/apme/59/11/JAMC-D-20-0013.1.xml

Doblas-Reyes, F.J., A.A. Sörensson, M. Almazroui, A. Dosio, W.J. Gutowski, R. Haarsma, R. Hamdi, B. Hewitson, W.-T. Kwon, B.L. Lamptey, D. Maraun, T.S. Stephenson, I. Takayabu, L. Terray, A. Turner, and Z. Zuo, 2021: Linking Global to Regional Climate Change Supplementary Material. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1363–1512, https://doi.org/10.1017/9781009157896.012

Doury A, Somot S, Gadat S, Ribes A, Corre L (2022). Regional Climate Model emulator based on deep learning: concept and first evaluation of a novel hybrid downscaling approach. Climate Dyn. https://doi.org/10.1007/s00382-022-06343-9

Giorgi F and Mearns LO (2002) Calculation of average, uncertainty range and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging". *Journal of Climate*, 15, 1141–1158. https://doi.org/10.1175/1520-0442(2002)015<1141:COAURA>2.0.20;2

Maraun D, Shepherd TG, Widmann M, Zappa G, Walton D, Gutiérrez JM, Hagemann S, Richter I, Soares PM, Hall A, et al. (2017): Towards process-informed bias correction of climate change simulations, Nature Climate Change, 7, 764. https://doi.org/10.1038/nclimate3418

Maraun D, Widmann M, Gutiérrez JM, Kotlarski S, Chandler RE, Hertig E, Wibig J, Huth R, Wilcke RAI (2015) VALUE—a framework to validate downscaling approaches for climate change studies. Earth's Future 3(1):1–14. https://doi.org/10.1002/2014EF000259

Maraun D, Widmann M, Gutiérrez JM (2019) Statistical downscaling skill under present climate conditions: A synthesis of the VALUE perfect predictor experiment. International Journal of Climatology, 39(9), 3692-3703. https://doi.org/10.1002/joc.5877

McGinnis S and Mearns L (2021) Building a climate service for North America based on the NA-CORDEX data archive, Climate Services, 22, 100233. https://doi.org/10.1016/j.cliser.2021.100233

Gutiérrez JM, Maraun D, Widmann M, et al. (2019) An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment. Int. J. Climatol.; 39: 3750– 3785. https://doi.org/10.1002/joc.5462

Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, and Prabhat (2019) Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, https://doi.org/10.1038/s41586-019-0912-1

Skific N, Francis JA, and Cassano JJ (2009) Attribution of Projected Changes in Atmospheric Moisture Transport in the Arctic: A Self-Organizing Map Perspective. J. Climate, 4135–4153, https://doi.org/10.1175/2009JCLI2645.1

Solman et al. (CORDEX SAT) 2021. The future scientific challenges for CORDEX. CORDEX web, https://cordex.org/wp-content/uploads/2021/05/The-future-of-CORDEX-MAY-17-2021.pdf

Vandal T, Kodra E, and Ganguly A R (2019) Intercomparison of machine learning methods for statistical downscaling: the case of daily and extreme precipitation, Theoretical and Applied Climatology, 137, 557–570, https://doi.org/10.1007/s00704-018-2613-3

Vaittinada P, Vrac M, Bastin S et al. (2016) Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: present climate evaluations. Clim Dyn 46, 1301–1329, https://doi.org/10.1007/s00382-015-2647-5

von Storch H and Zorita E (1999) The Analog Method as a Simple Statistical Downscaling Technique: Comparison with More Complicated Methods. J. Climate 12: 2474-2489.

Vrac M (2018) Multivariate bias adjustment of high-dimensional climate simulations: the Rank Resampling for Distributions and Dependences (R2D2) bias correction, Hydrol. Earth Syst. Sci., 22, 3175–3196,
https://doi.org/10.5194/hess-22-3175-2018

Xu Y, Gao X and Giorgi F (2010) Upgrades to the reliability ensemble averaging method for producing probabilistic climate-change projections. Climate Research, 41, 61– 81, https://doi.org/10.3354/cr00835