

Project: **ERA-5 Extract (E5E)**

Project title: **Konsortialdatenprojekt ERA5-Datenpool**

Project lead: **Michael Lautenschlager**

Allocation period: **2019-01-01 to 2022-12-31**

## Project overview – Abstract

In den letzten Jahren wurde von Wissenschaftlern vermehrt Interesse an hochauflösenden Reanalysedaten geäußert. Das gilt sowohl für Forscher aus dem Bereich der Wetter- und Klimamodellierung als auch für Beobachter, die ihre Daten an die Reanalysen anschließen wollen.

Aus dieser Nachfrage heraus wurden verschiedene grobe globale Reanalyse-Datensätze durch Downscaling bis zu Auflösungen von wenigen Kilometern verfeinert (COSMO-REA6, -REA2 u.a.). Diese Daten waren aber räumlich stets auf Europa, oft sogar nur auf Deutschland beschränkt.

Mit der neusten Version der ECMWF-Reanalysen (ERA-5) liegt erstmals ein globaler Datensatz hoher raumzeitlicher Auflösung vor (31km, 1h), der viele Jahrzehnte überdeckt (1950 bis zur Gegenwart). Neu gegenüber der Vorgängerversion ERA-Interim ist auch, dass ERA-5 EU-finanziert ist. Dadurch unterliegt sie den gleichen Lizenzrechten wie die Daten des Copernicus-Erdbeobachtungssystems und ist damit bei Ursprungsnennung völlig frei verfügbar. Im beantragten Projekt „ERA-5 Extract“ soll ein Auszug aus den vom ECMWF angebotenen ERA-5-Daten am DKRZ vorgehalten werden, der sich an den Datenbereichen orientiert, die bereits als ERA-Interim vom DKRZ angeboten werden. Andere Teile von ERA-5 (z.B. ERA-5-Land) bleiben soweit unberücksichtigt.

Darüber hinaus können die ERA-5-Daten Modellentwicklern als Referenz auf die Wetterentwicklung dienen, um die Qualität von Klimamodellen zu bewerten.

Der Zugriff auf diese großen Datenmengen ist direkt beim ECMWF möglich, jedoch sehr zeitaufwendig. Dies liegt auch an den Leitungsquerschnitten für das Datendownload, vor allem aber den langen Queueing-Zeiten bei Nutzung der Bandmaschinen des ECMWF, die man sich europaweit mit Wissenschaftlern und Vorhersageproduzenten teilen muss.

Aufbauend auf Nutzerbefragungen sowie auf den Erfahrungen aus der Verteilung von ERA-Interim-Daten schlagen wir zwei verschiedene Wege vor, die genannten Daten deutschen Forschern verfügbar zu machen – im zweiten Fall sind dabei dreiviertel der Daten (die Atmosphärendaten auf Modellleveln) nicht unmittelbar auf Festplatten verfügbar, sondern gecached auf Bänder ausgelagert. Sie führen zu unterschiedlichen Bedarfen an Speichermedien. Wir bitten den WLA beide Wege zu diskutieren und zu entscheiden.