Risk Modelling Technical Note

TA-9878 REG: Developing a Disaster Risk Transfer Facility in the Central Asia Regional Economic Cooperation Region

Contents

List of Abbreviations

ADB	Asian Development Bank
ADRC	Asian Disaster Reduction Center
ALCS	Afghanistan Living Conditions Survey
APHRODITE	Asian Precipitation – Highly-Resolved Data Integration Towards Evaluation of Extreme Events
CAREC	Central Asia Regional Economic Cooperation
CLIMADA	CLImate ADAptation risk modelling platform
CRED	Centre for Research on the Epidemiology of Disasters
CFR	Case Fatality Ratio
CMIP	Coupled Model Intercomparison Project
CORDEX	Coordinated Regional Climate Downscaling Experiment
COVID-19	Coronavirus disease
EMCA	Earthquake Model for Central Asia
EM-DAT	Emergency Events Database
EMME	Earthquake Model of the Middle East region
EPI	Epidemic Preparedness Index
GAR	Global Assessment Report on Disaster Risk Reduction
GCM	General Circulation Model
GEM	Global Earthquake Model
GFES	Global Flood Event Set
HAZ	Hydrological Accumulation Zones
IFRC	International Federation of Red Cross and Red Crescent Societies
IPCC	Intergovernmental Panel on Climate Change
JBA	Jeremy Benn Associates
MMI	Modified Mercalli Intensity Scale
NCEI	National Centers for Environmental Information
NEA	Northeastern Asia Earthquake Model
OCHA	United Nations Office for the Coordination of Humanitarian Affairs
PDO	Pacific Decadal Oscillation
PRC	People's Republic of China
PSHA	Probabilistic Seismic Hazard Analysis
RCM	Regional Climate Model
SEIR	Susceptible-Exposed-Infectious-Removed
SoP	Standard of Protection
TA	Technical Assistance
UN	United Nations

WCRP	World Climate Research Programme

List of Figures

List of Tables

About this document

The Asian Development Bank (ADB) is supporting CAREC member countries to strengthen their disaster risk management strategies and public sector budget resilience. This Technical Assistance (TA) forms part of such support. Specifically, this TA will profile earthquake, flood and infectious disease risk to inform the design of a regional risk transfer facility to spread and share such risk.

This document is a Technical Note that elaborates on the risk modeling undertaken as part of this TA. Several products and documents written under this TA rely on the data, modeling and research described in this Technical Note, namely:

- The **disaster risk profiles** which outline the hazard, exposure, and vulnerability characteristics of each of the eleven CAREC member countries. This Technical Note provides background on the data used, describes methodologies, and discusses limitations related to the data and the approaches.
- The **compound risk and infectious disease reports** which present infectious disease risk across the CAREC region and quantify the compounding of earthquake, flood, and infectious disease risk.
- The **protection gap assessment report** which discusses and estimates the ‘protection gap’ for flood and earthquake disaster risk in the CAREC region.
- The **Disaster Risk Management Interface (DRMI)** which has been designed to support knowledge development, awareness raising, and policy decision-making related to disaster risk across the CAREC region.
- The **Disaster Risk Management Interface User Guide** which provides an overview of the functions included in the DRMI.

Risk analysis

Definition of Risk Metrics and Terminology

The risk analysis section brings together exposure, hazard, and vulnerability inputs which are described in detail in dedicated sections of this technical note. Earthquake and flooding risk are described in the disaster risk profiles using a selection of standard risk metrics.

Affected People – the human impact of natural hazard events which may include disruption to livelihoods, income streams, injury, and fatalities.

Average Annual Fatalities (AAF) – the modelled fatalities resulting from flooding / earthquake shaking that is expected on average for a given year. Calculated at the country-level, and at the province level.

Average Annual Fatalities Ratio (AAFR) – the AAF normalized by the total population. The AAFR represents the proportion of the total population that is expected to be killed due to flooding / earthquake shaking. Calculated at the province level.

Average Annual Loss (AAL) – the modelled loss resulting from flooding / earthquake shaking that is expected on average for a given year. Calculated at the country-level, at the province level, and by asset type.

Average Annual Loss (Deaths) – the modelled fatalities resulting from infectious disease that is expected on average for a given year. Calculated at the country-level.

Average Annual Loss (Infections) – the modelled infections resulting from infectious disease that is expected on average for a given year. Calculated at the country-level, and at the province level.

Average Annual Loss Ratio (AALR) – the AAL normalized by the total exposed value of buildings. The AALR represents the proportion of the replacement value of the building stock that is expected to be lost due to flooding / earthquake damage. Calculated at the province level.

Average Annual Number of People Affected (AAPA) – the modelled number of people that are expected to be affected by flooding / earthquake shaking on average for a given year. Calculated at the country-level, and at the province level.

Direct Damages – losses that result from hazard event impacts to assets including infrastructure, buildings and their contents.

Disease cases exceedance probability curve – The modelled number of cases, of a given pathogen, that is expected on average for a given return period. Calculated for selected pathogens relevant to the country in question.

Earthquake / flooding damage exceedance probability curve – the modelled damages resulting from flooding / earthquake shaking, that is expected on average for a given return period. Calculated for direct and indirect damages.

Economic Damage and Loss – the direct and indirect damages and losses resulting from a natural hazard event. This includes impacts to the natural and built environment, to people, businesses, and governments.

Indirect Damages – losses that result from hazard event disruption to social, governance / administrative, and economic activities, critical and public services, and people's livelihoods.

Historical losses and impacts

Data sources and methodology

Several databases/sources were consulted to collate historical loss and impact data for each of the CAREC member countries, focusing on earthquake, flood, and infectious disease events since 1990. A description of each database is provided in *Table 1*.

Table 1: Referenced databases for collation of disaster impact and loss information across the CAREC member states

Database	**Description**
Asian Disaster Reduction Center	Established in 1988, the ADRC collates and archives key information on major disaster events from across the region. Detailed reports are additionally available for the largest-scale disaster events.
Emergency Events Database (EM-DAT)	The Centre for Research on the Epidemiology of Disasters (CRED) established EM-DAT in 1988. EM-DAT contains essential core data on the occurrence and effects of over 22,000 mass disaster events in the world from 1900 to present. EM-DAT is compiled from various sources, including UN agencies, non-governmental organizations, insurance companies, research institutes and press agencies. Historical coverage is variable by region, ranging from 1900 to present day. For a disaster event to be entered into EM-DAT, it must satisfy at least one of the following criteria: i) >=10 people reported killed, ii) >= 100 people reported affected, iii) declaration of a state of emergency, iv) call for international assistance.
FloodList	Established in 2008 to provide news and reports on major flooding events around the world. The site also acts as a historical database for news reports. FloodList is funded by the European Union Copernicus programme and is managed by a team of researchers based in Germany.
Global Significant Earthquake Database	Hosted by the National Centers for Environmental Information (NCEI), this databased provides a global listing of over 5,700 earthquakes from 2150 BC to the present. 'Significant' earthquakes are defined as those that caused i) moderate damage (approximately $1 million or more), ii) >= 10 deaths, iii) magnitude >= 7.5, iv) Modified Mercalli Intensity of X or greater, v) or the earthquake generated a tsunami.
International Federation of Red Cross and Red Crescent Societies (IFRC)	A global humanitarian organization, which coordinates and directs international assistance following natural and man-made disaster events in non-conflict situations. Produces reports covering the impacts of major disaster events.
Metabiota infectious disease database	A global database collated from >240 data sources, capturing >1,200 outbreaks over a 100-year period. The database covers 230 countries and territories and over 150 pathogens.
Munich Re	Munich Re's NatCatSERVICE has collated data on disaster losses since 1980. While the database itself is proprietary, relevant reports / articles were consulted to triangulate specific event impacts/losses.
National Centers for Environmental Information (NCEI)	An archive of atmospheric, coastal, geophysical, and oceanic research data, including dedicated section for Natural Hazards, Disasters and Severe Weather.

ReliefWeb	A humanitarian information service provided by the United Nations Office for the Coordination of Humanitarian Affairs (OCHA). Includes detailed articles and reports on major humanitarian disaster events.
Swiss Re sigma reports	Provides global assessments of insured and uninsured losses across a range of perils.
World Bank	Selected reports / materials consulted where triangulation of specific event impacts / losses was warranted.

Local data sources were also consulted and are referenced in the relevant section of each risk profile document. A list of countries for which local data sources were available is presented in Table 2. Loss data collated from these and selected local data sources was made consistent across profiles and over time by converting values to USD, consistent to 2019 prices. Impact and loss figures from these databases were validated through reference to national reports / sources where possible.

Table 2 Local data sources used to obtain historic loss information.

Country	**Data Source**
Azerbaijan	Ministry of Emergency Situations of The Republic of Azerbaijan. State Agency for Water Resources. Accessible at: https://www.fhn.gov.az/?eng
Kyrgyz Republic	Ministry of Emergency Situations of the Kyrgyz Republic. Accessible at: http://en.mes.kg/
Mongolia	Institute of Meteorology, Hydrology and Environment, Mongolia
Pakistan	National Disaster Management Authority of Pakistan. Accessible at: http://cms.ndma.gov.pk/
Tajikistan	Committee for Emergency Situations and Civil Defense, Government of the Republic of Tajikistan
Uzbekistan	Centre of Hydrometeorological Service under the Cabinet of Ministers of the Republic of Uzbekistan; The State Committee of the Republic of Uzbekistan on Statistics.

Uncertainties and limitations in the data and methodology

The historical loss information collected during this project represents an important input to risk modeling and helps to guide disaster risk management recommendations. Demonstrating the value of timely and accurate data collection following disaster events should provide a clear motivation for investment in appropriate data collection practices and in the institutions necessary to support them. As databases continue to be enhanced and updated, risk modeling outputs can be further refined, leading to improved model outputs.

Databases collating information on historical losses and impacts are characterized by certain uncertainties and limitations. These include:

- **Underreporting of losses and impacts.** This occurs for various reasons, including a lack of institutions responsible for collecting and archiving such information, lack of formal reporting criteria/changing criteria over time, difficulties surrounding the collation of information collected globally/in different languages;
- **Historical bias.** Loss and/or damage reporting completeness increases closer to present. This can result in the omission of smaller/moderate events further back in time, resulting in partial coverage that is bias towards the inclusion of larger events;

- **Local data sources.** Local sources of data were consulted where possible, though noting that sometimes such sources are difficult to access, collected on an ad hoc rather than structured basis, and more difficult to interpret in a consistent way.

The uncertainties and limitations of individual databases have been mediated to some extent by consulting a selection of databases, allowing for triangulation with complementary datasets (articles, industry and academic reports, peer-reviewed research papers). Additional checks have been undertaken where necessary through consultation with local stakeholders across the CAREC member states.

Hazard

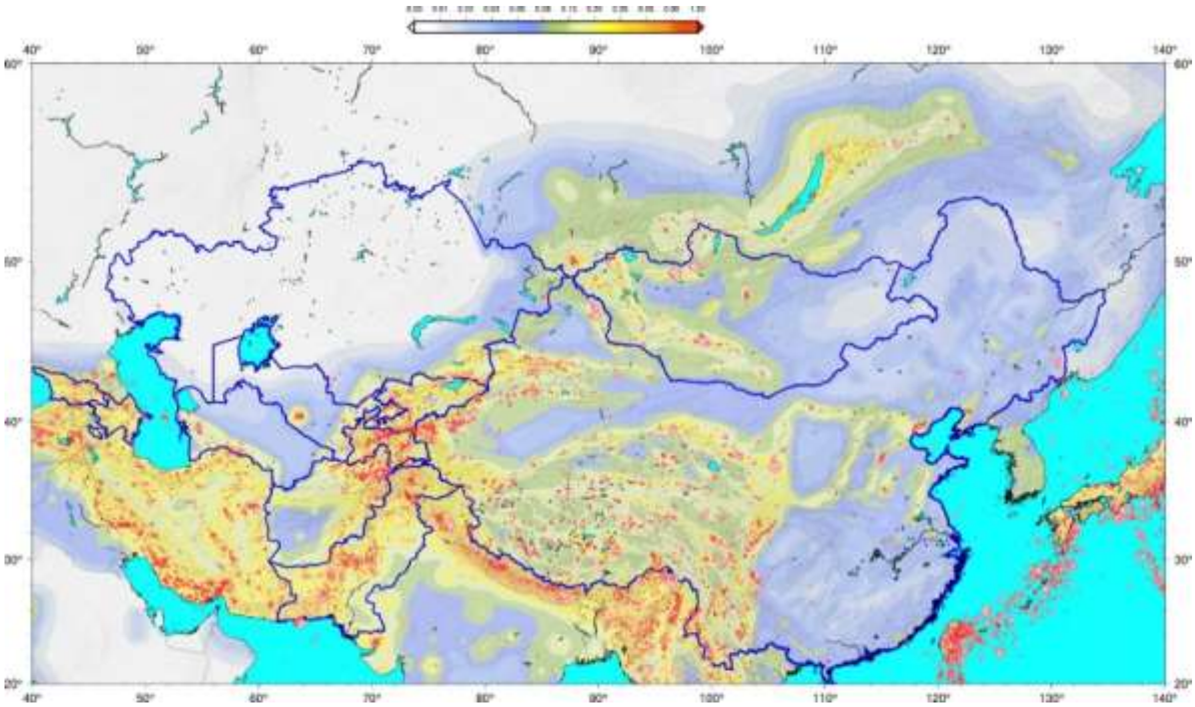
Data sources and methodology

Earthquake

Earthquake hazard was computed across the CAREC member countries using a selection of hazard models included in the GEM Hazard Mosaic.[1] The GEM Hazard Mosaic is a collection of the best earthquake hazard science available publicly at the global scale (Figure 1). The models used in this project are:[2] the Earthquake Model of the Middle East region[3] (EMME) covering the Middle East, the Earthquake Model for Central Asia[4] (EMCA) covering Central Asia and, the Northeastern Asia (NEA) model covering Mongolia.

The China Earthquake Administration did not authorize the use of the GEM Hazard Mosaic model for the People's Republic of China (PRC). GEM is currently developing a new model for PRC, with a first release version of this model used to compute hazard for the Inner Mongolia and Xinjiang Uyghur Autonomous Regions.

Figure 1: Seismic hazard map depicting the geographic distribution of the peak ground acceleration with 10% probability of exceedance in 50 years. Selected CAREC member countries are highlighted blue.



[1] Pagani M, Garcia-Pelaez J, Gee R, et al. (2020) The 2018 version of the Global Earthquake Model: Hazard component. *Earthquake Spectra,* 36(1_suppl), 226-251. https://10.1177/8755293020931866

[2] GEM Hazard Model Documentation. Accessible at: https://hazard.openquake.org/gem/

[3] Giardini, D., Danciu, L., Erdik, M., Sesetyan, K., Demircioglu, M., Akkar, S., Gülen, L. and Zare, M. (2018) Seismic Hazard Map of the Middle East. *Bulletin of Earthquake Engineering,* 16, 3567-3570. https://doi.org/10.1007/s10518-018-0347-3

[4] Ullah, S., D. Bindi, M. Pilz, L. Danciu, G. Weatherill, E. Zuccolo, A. Ischuk, N. N. Mikhailova, K. Abdrakhmatov, and S. Parolai. (2015) Probabilistic Seismic Hazard Assessment for Central Asia. *Annals of Geophysics,* 58 (1). https://doi.org/10.4401/ag-6687.

Seismic hazard maps were computed from the regional models using the OpenQuake Engine, the hazard and risk calculation engine developed by GEM[5,6], which follows the Probabilistic Seismic Hazard Analysis (PSHA) procedure.[7] PSHA involves the construction of an input model containing a seismic source and a ground-motion characterization. The input model defines the position of the earthquake sources, their geometry and the frequency with which each source generates events of different magnitudes. The latter specifies the ground-motion models used to compute the level of shaking at a collection of sites given a rupture.

The next step is the calculation of the so-called hazard integral. This integral is resolved numerically in the OpenQuake Engine using two strategies, one that samples ruptures and ground motion using a Monte Carlo procedure[8] and, one that solves the integral using discrete representations of functions. The former approach generates the information needed to compute risk while the latter calculates seismic hazard maps and hazard results used, for example, in building code regulations.

Flood

Flood risk modeling across the CAREC member countries has been conducted using JBA's probabilistic Global Flood Model at a consistent spatial resolution of 30 m across all countries / regions. The Global Flood Model incorporates JBA's Global Flood Map and Global Flood Event Set and is implemented using JBA's FLY technology. FLY is used to implement updated vulnerability (depth-damage) functions and disaggregate exposure values to coordinate points (weighted by WorldPop 2015 population distributions) across the area of interest. The result is a distribution of exposure points that is geographically continuous, allowing the full distribution of hazard intensities (i.e. flood depths) to be captured.

The Global Flood Event Set (GFES) is a catalogue of over 15 million plausible inland flood events worldwide. Events characterize the extent and intensity of flooding geographically and are represented by river and surface water severity at point locations. The GFES is generated using a modeling cascade which starts with the simulation of rainfall everywhere, then rainfall-runoff, and finally event selection. Rainfall simulation is performed using a unique combination of sophisticated statistical techniques, and rainfall-runoff is simulated everywhere using physically based hydrological models and prediction in ungauged basins. Importantly, the GFES contains flood events that are more extreme than have been observed in recent history, but which are still physically plausible.

A regionally calibrated rainfall-runoff approach allows the generation of river flooding at locations without flow gauges. The results of the simulations are time series of daily river flow and rainfall intensities which are then grouped into events using multivariate de-clustering.

The Global Flood Map provides undefended river and surface water flood extents and depths for six return periods (20, 50, 100, 200, 500 and 1,500 years). The map is created using observed river and rainfall data to generate extreme rainfall and river flow volumes and allowing those volumes to spread across the terrain using hydraulic modeling. River modeling captures flooding from rivers with a catchment area over 500 km^2 and surface water modeling

---

[5] Pagani, M., D. Monelli, G. Weatherill, L. Danciu, H. Crowley, V. Silva, P. Henshaw, et al. (2014) OpenQuake Engine: An Open Hazard (and Risk) Software for the Global Earthquake Model. *Seismological Research Letters*, 85(3), 692–702. https://doi.org/10.1785/0220130087.

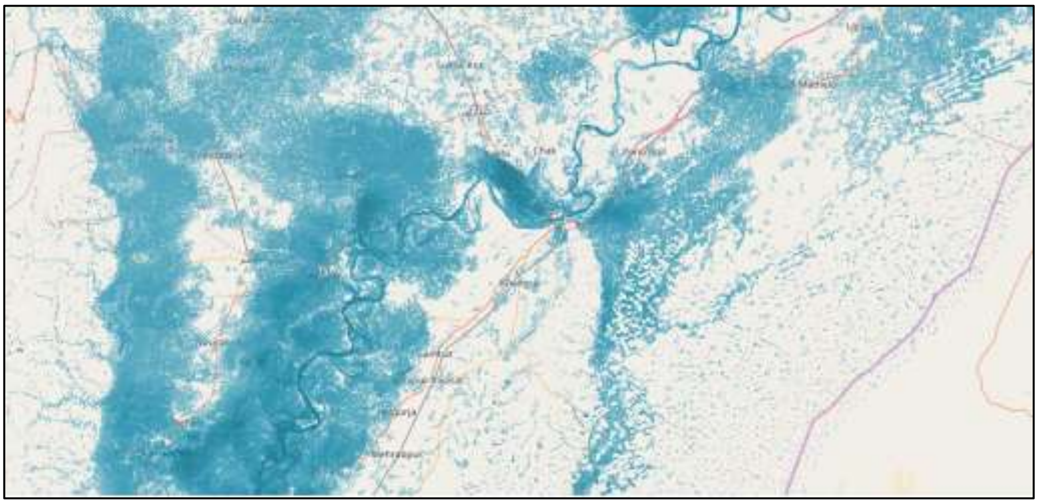
[6] Silva, V., Crowley, H., Pagani, M., Monelli, D. and Pinho, R. (2014) Development of the OpenQuake Engine, the Global Earthquake Model's Open-Source Software for Seismic Risk Assessment. *Natural Hazards,* 72(3), 1409–27. https://doi.org/10.1007/s11069-013-0618-x.

[7] Cornell, C.A. (1968) Engineering seismic risk analysis. *Bulletin of the Seismological Society of America,* 58(5), 1583–1606. https://doi.org/10.1785/BSSA0580051583

[8] A Monte Carlo sampling approach involves random sampling of a probability distribution (in this case the distribution of rupture locations and ground motion intensity) to capture the plausible range of hazard events.

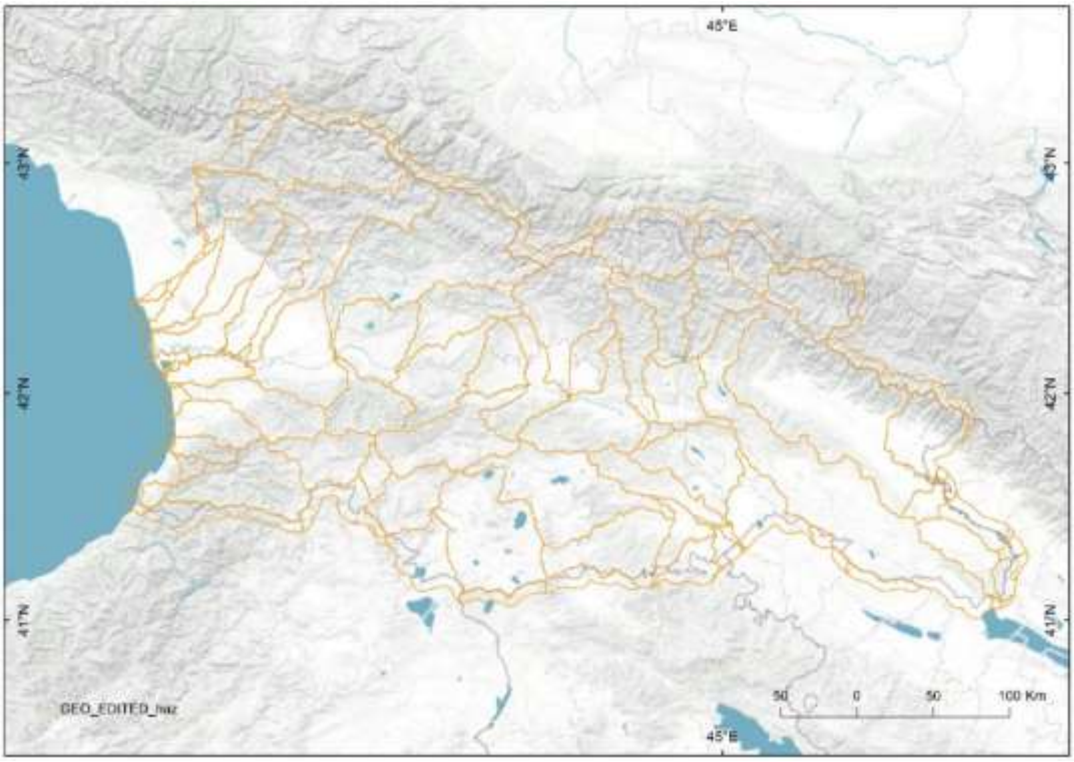
captures flooding from smaller rivers and locations where rainfall pools in depressions in the topography (*Figure 2*).

Figure 2: Example of a 100-year return period Global Flood Map for a small area of the Indus basin in Pakistan



The river flood areas used in the Global Flood Model are hydrologically sensible areas known as Hydrological Accumulation Zones (HAZ). HAZ provide a simple means of identifying areas that may be affected by the same flood event and provide a consistent approach at national or global scale (**Error! Reference source not found.**). Each HAZ represents the boundary of a hydrological catchment approximately 500 km^2 in size and are generated using Digital Elevation Models and drainage lines.

Figure 3: Example of the hydrological catchments used for flood modeling.



Flood defences are modelled in parallel to JBA's undefended flood maps and the two datasets are combined to produce a defended view of flooding. The defence locations and extents are determined using a combination of third-party sources (e.g. published government information, aerial imagery, engineering reports, other online material). Where third-party information is of poor quality or incomplete, JBA's own analysis is used to determine the defence location and the likelihood that the defence measure will be rendered ineffective due to flooding from undefended areas. Defences included in Global Flood Model include permanent, physical barriers (e.g. dykes) and built defence schemes. Temporary or demountable defences are not included.

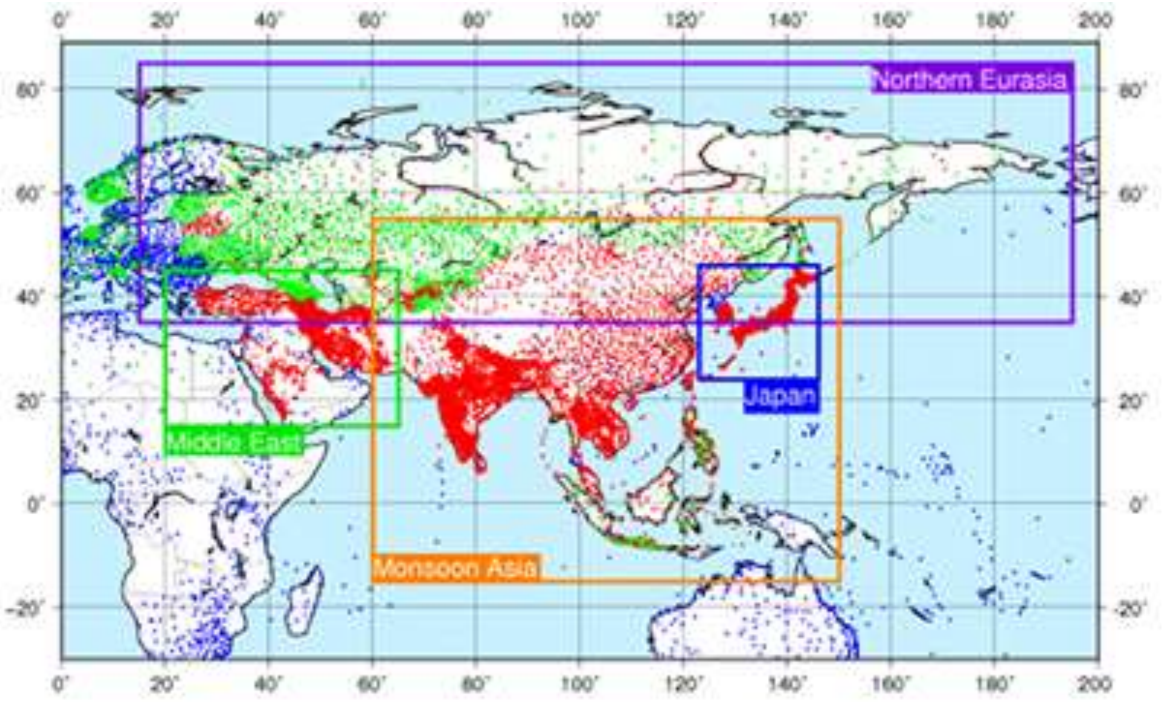
The area that benefits from protection is identified for each flood defence installation and a Standard of Protection (SoP), expressed as a return period, is then attributed to the defended area. Where the severity (as a return period) of a simulated flood event in Global Flood Model remains below the SoP, the defence is modelled as being fully effective and no flooding occurs. Where the severity of a simulated flood event exceeds the capacity of the defence, a defence overtopping calculation is applied to reduce the impact of the flooding based on the volume of water overtopping the defence. The change in flood return period is applied to all exposure points located within a defended area.

Historical Climate Data

Historical climate data was obtained from regional climate models, bias corrected using the Asian Precipitation – Highly-Resolved Data Integration Towards Evaluation of Extreme Events (APHRODITE) dataset.[9] This dataset contains the greatest number of valid stations for the continent in comparison with other available gridded climate datasets, due to the cooperative agreements with national agencies to secure data (Figure 4).

The APHRODITE gridded daily precipitation data is a 50+ year dataset (derived from observational gauge and remote sensing data) covering Monsoon Asia, Central and East Asia, the Middle East and parts of the Russian Federation at a grid resolution of 0.25°x0.25°. For the Monsoon Asia domain, encompassing this project's countries Pakistan, Tajikistan and Afghanistan, the dataset covers the period 1951-2015. For all other project countries, APHRODITE data is available for the period 1951-2007.

Figure 4: APHRODITE rain gauge distributions for the different data domains



Source: http://aphrodite.st.hirosaki-u.ac.jp/products.html

Future Climate Projections

Projections of future precipitation under the representative concentration pathways (RCP) RCP4.5 and RCP8.5 were obtained from the Coordinated Regional Climate Downscaling Experiment (CORDEX).[10] Established by the World Climate Research Program (WCRP) in 2009, CORDEX is an international cooperative climate modeling experiment, through which multiple Regional Climate Models (RCMs) are driven by general circulation models (GCMs) from the Coupled Model Intercomparison Project 5 and 6 (CMIP5 and CMIP6) - the simulations of which inform the Intergovernmental Panel on Climate Change (IPCC) reports. For this project, the following multi-model precipitation projections were used to develop intensity-

[9] Yatagai, A., K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi and A. Kitoh (2012) APHRODITE: Constructing a Long-Term Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. *Bulletin of the American Meteorological Society*, 1401-1415. https://doi.org/10.1175/BAMS-D-11-00122.1

[10] Giorgi, F., Jones, C. and Asrar, G. (2009) Addressing climate information needs at the regional level: the CORDEX framework. *WMO Bulletin*, 58(3), 174-183.

duration-frequency curves of potential future extreme rainfall events due to climate change (Table 3).

Table 3: List of GCM-RCM models from the CORDEX initiative used in this project.

CORDEX South Asia Domain – used for Pakistan, Tajikistan and Afghanistan		
Experiment Name	**RCM Description**	**Driving GCM**
CCCma-CanESM2-IITM-RegCM4	The Abdus Salam International Center for Theoretical Physics (ICTP) Regional Climate Model version 4.4.5 (RegCM4)	Canadian Centre for Climate Modelling and Analysis (CCCma) - second generation Canadian Earth System Model (CanESM2)
CSIRO-Mk3.6-IITM-RegCM4		Commonwealth Scientific and Industrial Research Organization (CSIRO) - Mk3.6
IPSL-CM5A-LR-IITM-RegCM4		Institut Pierre-Simon Laplace (IPSL) - CM5A Earth System Model
MPI-ESM-MR-IITM-RegCM4		Max Plank Institute for Meteorology (MPI) - Earth System Model
NOAA-GFDL-ESM2M-IITM-RegCM4		National Oceanic and Atmospheric Administration (NOAA), Geophysical Fluid Dynamics Laboratory (GFDL) - Earth System Model (ESM2M)
CNRM-CERFACS-CM5-SMHI-RCA4	Rossby Centre, Swedish Meteorological and Hydrological Institute (SMHI) - regional atmospheric model version 4	National Centre for Meteorological Research (CNRM) and Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS) - Earth System Model
CORDEX East Asia Domain: Mongolia and PRC (Inner Mongolia and Xinjiang Uygur Autonomous Regions)		
CNRM-CERFACS-CM5-CLMcom-CCLM5		CNRM and CERFACS
IHECH-EC-EARTH-DMI-HIRHAM5	Danish Meteorological Institute (DMI) - regional hydrostatic climate model (HIRHAM5)	Irish Centre for High-End Computing (ICHEC), European Consortium (EC) - Earth System Model
IHECH-EC-EARTH-CLMcom-CCLM5		ICHEC-EC
MPI-ESM-LR-CLMcom-CCLM5		MPI
CORDEX Central Asia Domain: Azerbaijan, Georgia, Kazakhstan, Kyrgyz Republic, Turkmenistan, and Uzbekistan		
MOHC-HadGEM2-ES-BOUN-RegCM4		
MPI-ESM-MR-BOUN-RegCM4		MPI

Daily gridded precipitation data were downloaded for each experiment from an Earth System Grid Federation data node. For historical simulations, data was obtained for the period 1961-2005 (CORDEX SA and EA); and 1970-2005 (CORDEX CA). For future simulations, data was obtained for two RCPs (4.5 and 8.5) for the period ~2031-2070 (centered around the 2050s).

A 45-year window was selected over the standard 30-year climatology window for the CORDEX South Asia and East Asia domains, as these monsoon regions are influenced by multi-decadal teleconnections such as the Pacific Decadal Oscillation (PDO).[11,12] The PDO undergoes phase shifts approximately every 20 to 30 years, requiring longer timeseries (ideally 50+ years)[13] to capture such shifts and their influences on regional Asian climates. A 36-year window was applied to the CORDEX Central Asia domain, as the historical GCM-RCM model simulations were only available for the period 1970-2005.

While RCMs are able simulate localized climate features much better than GCMs, they still often have significant biases.[14] Precipitation biases in historical simulations can be quite large for precipitation extremes in particular;[15] these biases may get carried forward into future projections and even be further inflated. Due to these biases, it is not advised that RCM projection data be directly used for climate change impact assessment studies, such as for future flood model estimation, without prior bias correction.

For this analysis, bias correction was undertaken using a quantile mapping technique. Quantile mapping involves developing a transfer function comparing the distributions of the observed historical data with modeled historical simulations and transforming the distribution of the modeled variable to match the distribution of the historical.[16]

This transfer function is then applied to future model projections.[17] However, RCM biases might not be stationary. That is, biases over the historical period might not persist exactly in the same way in the future. At the same time, it is likely that the real climate change signal may shift a variable's cumulative distribution function in the future. A method, quantile delta mapping, corrects biases from the historical period while preserving future projected relative changes in variable quantiles.[18] The bias corrected future projections are then found by multiplying the relative change function with the historical modeled bias corrected value.

A number of parametric and nonparametric transfer functions have been used in different studies. After testing the performance of a subset of parametric transfer functions, we found using the empirical cumulative distribution functions (a nonparametric method) to adjust the modeled values performed better and required less computational resources for the test country Pakistan. Pakistan was chosen as the test country given the complexity of its topography and the influence of the South Asian Monsoon system and westerlies in various provinces on seasonal precipitation. We used then this method for the remaining countries.

---

[11] Fu, C., Z. Jiang, et al. (2008) Chapter 3: Interdecadal Climate Variability in China Associated with the Pacific Decadal Oscillation in: Regional Climate Studies of China, pp. 97-117. Yang, X. and Y. Zhu [eds.], Springer Nature: Switzerland.

[12] Krishnan, R. and M. Sugi (2003) Pacific decadal oscillation and variability of the Indian summer monsoon rainfall, *Climate Dynamics*, 21, 233-242. https://doi.org/10.1007/s00382-003-0330-8

[13] Deser, C., K. Trenberth and NCAR (2016) "The Climate Data Guide: Pacific Decadal Oscillation (PDO): Definition and Indices." Retrieved from https://climatedataguide.ucar.edu/climate-data/pacific-decadal-oscillation-pdo-definition-and-indices.

[14] Kjellström, E., F. Boberg, M. Castro, et al. (2010) Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models. 4(2-3), 135-150, *Climate Research*, https://doi.org/10.3354/cr00932

[15] Christensen, J., F. Boberg, B. Christensen, et al. (2008) On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophysical Research Letters*, 35(20) https://doi.org/10.1029/2008GL035694.

[16] Dosio, A. and P. Paruolo (2011) Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: Evaluation on the present climate. *Journal of Geophysical Research*, 116(D16). https://doi.org/10.1029/2011JD015934

[17] Li, H., J. Sheffield and E. Wood (2010) Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. 115(D10). *Journal of Geophysical Research*, https://doi.org/10.1029/2009JD012882

[18] Cannon, A., S. Sobie and T. Murdock (2015) Bias Correction of GCM Precipitation Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate*, 28(17), 6938-6959. https://doi.org/10.1175/JCLI-D-14-00754.1

Finally, the intensity-duration-frequency shifts in future (2050s) precipitation extremes due to climate change (RCP4.5 and RCP8.5) were calculated assuming an Extreme Value Type I probability distribution function.[19] The approach to calculating the 24-hr maximum precipitation intensity shifts in the 2050s due to climate change can be summarized as:

- *Step 1*: Regridding the CORDEX data to match the coordinate system of APHRODITE. The CORDEX data are in a rotate polar coordinate system, while the APHRODITE data are in a regular geographic longitude/latitude grid. The CORDEX data were regridded using Climate Data Operators (CDO).

- *Step 2*: APHRODITE and the regridded CORDEX data were extracted for the Administrative Level 1 for each country.

- *Step 3*: Each month was extracted, and quantile delta mapping performed separately per month. Months were modeled separately to account for stark seasonal differences in precipitation amounts and dry day sequences.

- *Step 4*: The monthly bias corrected future projections were then recombined into annual datasets and the maximum annual precipitation value extracted for each year.

- *Step 5*: The maximum annual 24-hr projected rainfall values from the 25th, median and 75th quartiles projection member across the model suite were then used to calculate the 24-hr intensity-duration-frequency curves for return periods of: 2, 5, 10, 20, 50, 100, 200, 500, 1000, 1500, 5000 and 10,000 years.

Infectious Disease

Infectious disease risk profiles were modelled for each of the CAREC member states, individually and as a region. The risk profiles include probabilistic infection, hospitalization, and mortality risk profiles for each CAREC member state, broken down by pathogen. This work also forms the basis for modelling the potential trajectory of the COVID-19 pandemic within the region, as well as the impact of a potential clash between a natural hazard and an epidemic or pandemic, and risk financing best practices and options.

Infectious disease risk profile modeling

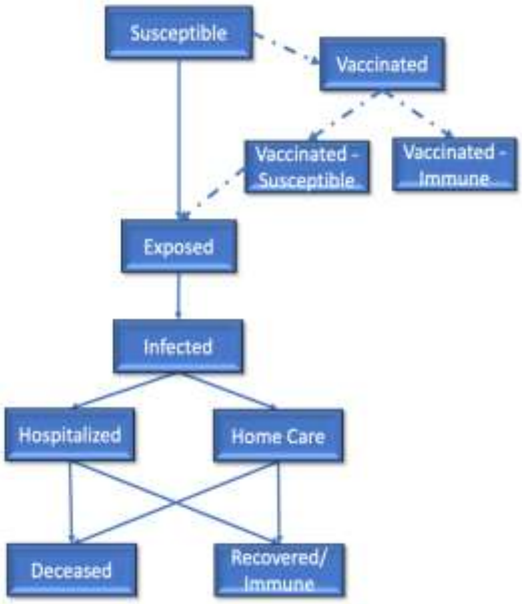
While the infectious disease modeling also captures elements of hazard, exposure, and vulnerability, the overall framework differs from that used for earthquake and flood. The infectious disease model uses a pathogen-specific, compartment model framework (see Figure 5 for an example), which incorporates the epidemiological dynamics of the disease of interest, including the availability of pharmaceutical responses (e.g., vaccines.)[20]

The disease spread model operates over a human population layer that integrates the population and demographic characteristics of each subpopulation. Two types of human mobility patterns are explicitly modeled: long-range mobility, which happens on a daily time scale, and short-range mobility (commuting) which happens over a time period shorter than one day. The modeled world is divided into 1,413 geographic regions; each considered to be a subpopulation. CAREC countries are represented across 52 geographical areas. Thus,

[19] Chow, V., D. Maidment and L. Mays (1988) *Applied Hydrology*. Civil Engineering Series. pp: 572. McGraw-Hill Book Company: New York.
[20] Madhav, N., Stephenson, N., Oppenheim, B (2021). Multi-Pathogen Event Catalogs: Technical Note. Washington, D.C.: World Bank Group. Accessed at: http://documents.worldbank.org/curated/en/181791625232959415/Multi-Pathogen-Event-Catalogs-Technical-Note

inputs / assumptions and outputs were made for each country individually, but also across different sub-groups and the entire CAREC regional bloc.

Figure 5: Example of the pandemic influenza compartmental model structure



Dividing countries into subpopulations allows for specificity in modeling epidemic and pandemic outbreaks. Within each subpopulation, disease accumulates during an outbreak simulation and may spread to neighboring subpopulations (via the short-range mobility component) or far-reaching subpopulations (via the long-range mobility component).

The disease spread model simulates the progression of each modeled event on a daily timestep, including population mobility, infections, hospitalizations, and deaths. This creates time series data for each modeled event.

Outbreak preparedness varies on an individual country basis. Preparedness can be defined as capacity to detect and respond to outbreaks. Metabiota has constructed a country-specific Epidemic Preparedness Index (EPI),[21] an estimator of epidemic preparedness covering 188 countries globally (including all CAREC member states). The EPI measures national capacity to effectively detect, report, and respond to public health emergencies, which will impact the regional occurrence and spread of modeled events. The EPI methodology, data, and back-testing and validation results have been published in a peer-reviewed scientific journal, and have been cited and used in the scientific literature, as well as in analyses of global pandemic preparedness as well as macroeconomic analyses conducted by the World Bank.

The model incorporates country-specific preparedness via the EPI, rather than using aggregated, regional assumptions. Variation in capacity impacts the regional occurrence and spread of modeled events by affecting the time until detection, the scope and scale of response mobilization, and the case counts and temporal dynamics of an epidemic simulation.

The modeled exceedance probability curves include only those infections and deaths that are in excess of the regularly occurring annual baseline. For the included respiratory diseases like pandemic influenza and novel coronaviruses, this baseline will be zero, but for diseases like Crimean-Congo Hemorrhagic Fever, which is endemic in some CAREC countries, the baseline will be higher than zero.

[21] Oppenheim B, Gallivan M, Madhav NK, et al. (2019) Assessing global preparedness for the next pandemic: development and application of an Epidemic Preparedness Index, *BMJ Global Health*; 4:e001157.

COVID-19 Scenario modeling

For this TA, Metabiota provided simulations of the potential trajectory of the COVID-19 pandemic, using a customized compartment model framework that incorporates the epidemiological dynamics of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2, the etiological agent of COVID-19). This computational model was used to simulate the spatiotemporal dynamics of SARS-CoV-2 within CAREC members. The model incorporates both potential pharmaceutical and non-pharmaceutical interventions – for example, the availability of vaccines, or implementation of social distancing measures – which are implemented differently in each country, based upon the degree of implementation of appropriate policy responses and mitigation measures.

Infectious disease modeling for compound risk analysis

The infectious disease natural hazard compounding modeling relies on the compartmental model structure described above, input datasets relevant to each CAREC member country / region, and associated assumptions surrounding intervention.

For the compound modeling specifically, baseline assumptions and parameter values were selected to reflect a pandemic influenza event with a return period of approximately 100 to 200 years. Simulations were initiated with 100 initial infections and run for three years. The modeled scenarios included vaccination of the susceptible population starting at 9 months, accounting for development time of a vaccine for a novel pandemic influenza virus which is applied for all CAREC countries.

Once vaccination begins in the scenario, the susceptible population is vaccinated at a constant rate, which varies by epidemic preparedness until 65% of the population is vaccinated or until the simulation ends. The vaccine efficacy within the simulation is assumed to be 75%. The baseline reproduction number is 1.9 and decreases over time to 1.5 due to non-pharmaceutical interventions (e.g., social distancing, etc.). The case fatality ratio (CFR) varies by epidemic preparedness between 0.003 and 0.006. The baseline parameter value assumptions are consistent with parameter values observed during prior influenza pandemics.

Uncertainties and limitations in the data and methodology

Earthquake

The four hazard input models developed use different dataset and methodologies for the eleven CAREC countries. This means that some differences in the hazard results at borders between models are inevitably present. Overall, these differences can be considered acceptable given the different approaches and components included in the different models (e.g. ground motion models). It is worth noting that the earthquake risk analyses were completed at the national scale, hence the model building approaches used for the various models did not impact on the internal consistency of the results computed in each country.

The models used for this work are regional models covering large continental areas which means that they do not necessarily reach the level of detail expected for a hazard model at the national scale. Moreover, model evaluation and testing activities carried out recently emphasized the need of improving some of models utilized in this project (e.g., Earthquake Model for Central Asia).

Flood

Flood is a complex natural hazard that can have impacts on a wide range of spatial scales. While the physical processes that result in river flooding (also known as fluvial flooding) are relatively well understood and can be effectively modelled, surface water flooding, also referred to as pluvial or flash flooding, presents a challenge. Furthermore, the locations where surface water flooding occurs rarely have rainfall measurements or monitoring gauges. This means that there is little historical information on which to base the models. Limitations in computing power means that it is difficult to model surface water flooding on the spatial scales on which it occurs. Where surface water occurs in urban areas, the location of property (including unplanned settlements) and the status of urban drainage networks has a significant impact on the flood risk. In mountainous parts of the region, snowmelt and glacial outflows can also have an impact on flood risk.

The Global Flood Model was regionally validated using observations from across the CAREC member countries. This procedure is impacted by the density of observation points. For Central Asia the spatial availability of gauge data is limited, leading to uncertainty in the application of gauged catchment characteristics to distant ungauged basins.

The intensity of flooding varies greatly over small spatial scales, so the resolution at which an analysis is carried out can have a considerable impact on estimated losses. For instance, the modeling of Hydrological Accumulation Zones (HAZ) depends on analysis of Digital Elevation Models whose resolution and accuracy is globally variable.

Historical Climate Data

Although APHRODITE remains the only long-term (1951 onward) high-resolution precipitation and temperature dataset for Asia, there are uncertainties associated with its use. For example, the density of observation stations declined in some countries in the 1990s post-Soviet Union, potentially resulting in an over- or under-estimation of precipitation (or temperature) in certain areas at a particular time period in the APHRODITE data.

Topography also introduces data uncertainties. Precipitation interpolation at high mountain elevations of the Hindu Kush Himalaya can be uncertain due to the combination of low spatial density of ground observations and micro-orographic effects. Complex topographies lend to significant micro-orographic effects on precipitation, with some studies,[22,23] noting a factor of four difference in precipitation amounts over less than 10 km in some parts of the Himalaya. These effects introduce uncertainties in precipitation amounts in high mountain areas, regardless of whether the precipitation is associated with the Asian Monsoon System, winter westerlies flow from the Mediterranean or more localized convection.

Yet, the quality of the APHRODITE datasets is strengthened in nearer decades through the incorporation of remote sensing observations and improved topography calculations. The underlying observational data undergo stringent quality control and cleaning before incorporation into the interpolation model. For these reasons, the APHRODITE dataset remains the preferred high-resolution, long-term gridded climate dataset for Asia.

Future Climate Projections

Multiple sources of uncertainty influence climate projections. Some uncertainty originates from within the models themselves, owing to imperfect understanding of the Earth system. Furthermore, certain models are designed to resolve some processes (e.g. convection) at higher resolution or for particular areas better than others.[24] The sources of uncertainty internal to climate models are being reduced with better understanding of land-ocean-atmospheric dynamics and improved computational power allowing for more model realizations to find and reduce model uncertainties.

Other sources of uncertainty are external to the climate models, for example, while it is well-established that land use change and emissions are driving climate change, it is difficult to precisely quantify how severe both will be in the future. The current RCPs used to drive models are plausible estimates of different climate futures, conditioned on scenarios of emissions trajectories. How much emissions and land use change continue to increase and at what rate, depends significantly on the policies and actions societies take to transform economies to be low carbon and sustainable.

For these reasons, it is necessary to use multiple climate models to derive a range of possible change in an area's climate. A larger number of models can provide a more robust range of future climate change (based on existing knowledge) than a single model or a few.

In this study, only two GCM-RCM combinations were available for the CORDEX Central Asia domain at the start of the project for which scientific literature evaluating model performance existed. And, as described previously, there are limitations with observational data in many countries. Over the South Asia domain, the combination included 6 GCMs, but only 2 RCMs. For the East Asia domain, results from 3 GCMs and 2 RCMs were available. Therefore, for all three regions the potential range of changes in precipitation extremes and seasonal means might not be as robust. Repeating the exercise for all three regions is advised once new GCM-RCM combinations become available in the future.

Infectious Disease

Metabiota's disease spread models include several parameters related to response, containment measures, and interventions. These are all included as components within the mechanistic model. For each pathogen catalog, the parameter distributions (e.g. timing of

[22] Lang, T. and A. Barros (2002) An investigation of the onsets of the 1999 and 2000 monsoons in central Nepal. *Monthly Weather Review*, 130(5), 1299-1316. https://doi.org/10.1175/1520-0493(2002)130<1299:AIOTOO>2.0.CO;2.

[23] Barros, A., G. Kim, E. Williams and S. Nesbitt (2004) Probing orographic controls in the Himalayas during the monsoon using satellite imagery. *Natural Hazards and Earth System Sciences*, 4, 29-51. https://doi.org/10.5194/nhess-4-29-2004.

[24] Giorgi, F., Jones, C. and Asrar, G. (2009) Addressing climate information needs at the regional level: the CORDEX framework. *WMO Bulletin*, 58(3), 174-183.

vaccine availability, efficacy of non-pharmaceutical interventions, etc.) that are included are based on the best available data and assessments of current outbreak response capabilities. Model assumptions regarding the variation in transmission rates and the effectiveness and timing of mitigation measures were assessed through sensitivity testing.

The impacts of potential improvements in containment efforts (e.g., faster response time, availability of a new vaccine) were tested within the modeling framework by modifying the assumptions surrounding these parameter distributions and would be reflected in the modeled outcomes.

Exposure

Data sources and methodology

A consistent exposure dataset was used across the earthquake and flood risk modeling, based on the GEM Global Exposure Database.[25] The exposure database contains information regarding the number of buildings, geographical location, replacement costs (including the structural and nonstructural components, and the building contents), number of occupants and vulnerability classes of the building stock. The GEM Building Taxonomy (Version 2.0)[26] was used to classify the building stock in the CAREC member countries.

We are leveraging the regional and national exposure models that formed the basis of GEM's global seismic risk modeling effort.[27] Four exposure models cover the region of interest for this project: the exposure model for Central Asia covering Kazakhstan, the Kyrgyz Republic, Tajikistan, Turkmenistan, and Uzbekistan; the exposure model for PRC; the exposure model for the Middle East covering Afghanistan, Azerbaijan, Georgia, and Pakistan; and the exposure model for Mongolia.

The German Research Centre for Geosciences led the development of the harmonized building exposure model for Central Asia,[28,29] and these models have been subsequently improved by GEM. The exposure model for the Xinjiang Uyghur and Inner Mongolia Autonomous Regions of the PRC derive from the national exposure model for the PRC developed by GEM in collaboration with Beijing Normal University, relying on a compilation of housing, establishment, and industry datasets obtained from the National Bureau of Statistics of China.[30] Exposure models for Afghanistan, Azerbaijan, Georgia, Pakistan, and Mongolia have been developed by GEM, based on the respective national population and housing censuses and establishment / enterprise surveys.

Overall, the exposure data includes residential, commercial, and industrial built assets for all member countries. The exposure models for each country are defined at the smallest administrative level at which source datasets are available. The development of the exposure models in all cases followed four main steps:

1. Definition of building classes.
2. Mapping census data to building classes.
3. Mapping housing units or establishments to buildings.
4. Estimation of built up areas and replacement costs.

[25] GEM Global Exposure Database. Accessible at: https://storage.globalquakemodel.org/what/physical-integrated-risk/exposure-database/

[26] Brzev, Scawthorn, C, Charleson, E.W., Allen, L., Greene, M., Jaiswal, K., and Silva, V. (2013) GEM Building Taxonomy (Version 2.0). *GEM Technical Report 2013-02.* Accessible at: https://pubs.er.usgs.gov/publication/70058718

[27] Silva, V., D. Amo-Oduro, A. Calderon, C. Costa, J. Dabbeek, V. Despotaki, L. Martins, et al. (2020) Development of a Global Seismic Risk Model. *Earthquake Spectra*, February. https://doi.org/10.1177/8755293019899953.

[28] Wieland, Marc, Massimiliano Pittore, Stefano Parolai, Ulugbek Begaliev, Pulat Yasunov, Sergey Tyagunov, Bolot Moldobekov, Saidislom Saidiy, Indalip Ilyasov, and Tanatkan Abakanov. "A Multiscale Exposure Model for Seismic Risk Assessment in Central Asia." Seismological Research Letters 86, no. 1 (2015): 210–22. https://doi.org/10.1785/0220140130.

[29] Wieland, Marc, Massimiliano Pittore, Stefano Parolai, Ulugbek Begaliev, Pulat Yasunov, Jafar Niyazov, Sergey Tyagunov, et al. "Towards a Cross-Border Exposure Model for the Earthquake Model Central Asia." Annals of Geophysics 58, no. 1 (2015). https://doi.org/10.4401/ag-6663.

[30] Ma, J., Rao, A., Silva, V. et al. (2021) A township-level exposure model of residential buildings for mainland China. *Nat Hazards,* 147. https://doi.org/10.1007/s11069-021-04689-7

The following sections of this document briefly describe these steps; and a more detailed overview of the general methodology adopted for the development of the exposure models can be found in Yepes-Estrada et al.[31]

Definition of building classes

The first step in developing an exposure model is the definition of classes that allow the grouping of buildings with similar characteristics. The current model includes, at least, information regarding the material and type of the lateral load resisting system, number of storeys, ductility level, and occupancy class. There are other attributes that might affect the buildings' structural response to damage, but most of the countries have limited information about those properties, and only a handful reported additional features that were included in the models.

The identification of the predominant building classes in each country depends on the construction practice, meteorological conditions and availability of construction materials. Local, national, and regional reports about predominant construction classes and damage surveys of past earthquakes in each country have been consulted, along with additional information from global initiatives, such as the World Housing Encyclopedia, UN-HABITAT, and the PAGER building inventory database. For Afghanistan and Pakistan, the list of building classes was also informed by the building typology classification for South Asia proposed in a joint study by the Aga Khan Development Network (AKDN) and the Norwegian Seismic Array (NORSAR)[32].All buildings in the overall exposure model for the CAREC member countries have been classified according to an extended version of the GEM building taxonomy. An example of the taxonomy applied to a selection of the CAREC member countries is provided in Table 4.

[31] Yepes-Estrada, Catalina, Vitor Silva, Anirudh Rao, Alejandro Calderón, Catarina Costa, Jamal Dabbeek, Luís Martins, Ana Beatriz Acevedo, Helen Crowley, and Murray Journeay. (2020) Development of a Global Exposure Model for Regional Seismic Risk Assessment. In *17th World Conference on Earthquake Engineering,* Sendai, Japan, 2020.

[32] Lang, D. H., Kumar, A., Sulaymanov, S., & Meslem, A. (2018). Building typology classification and earthquake vulnerability scale of Central and South Asian building stock. Journal of Building Engineering, 15, 261–277.

Table 4: Building classification used in the exposure models for Kazakhstan, the Kyrgyz Republic, Tajikistan, Turkmenistan, and Uzbekistan

Building Class	Description	Taxonomy (GEM v2.0)	Number of Occupants (Night)	Average Floor Area per Unit (m²)	Average Cost per Unit Area (USD/m²)	Average Cost per Unit (USD)
URM1	unreinforced masonry	/MUR+CLBRS+MOC/LWAL+DNO/FW	3.8	488	$ 175	$ 85,450
URM2	unreinforced masonry, concrete floors	MUR+MOCL/LWAL+DNO/FC	3.8	493	$ 200	$ 98,669
CM	confined masonry	/MCF+MOC/LWAL+DNO/FC/HBET	76	2,322	$ 288	$ 668,736
RM-L	reinforced masonry, low rise	/MR+MOC/LWAL+DNO/FC/HBET:1	5.2	2,054	$ 160	$ 359,488
RM-M	reinforced masonry, medium rise	/MR+MOC/LWAL+DNO/FC/HBET:3	104	3,380	$ 160	$ 540,731
RC1	RC frame without ERD	/CR+CIP/LFM+DUC/FC/HBET	152	4,208	$ 425	$ 359,488
RC2	RC frame with moderate ERD	/CR+CIP/LDUAL+DNO/FC/HBET:7	152	18,412	$ 425	$ 7,825,189
RC3	RC frame with high level of ERD	/CR+CIP/LFINF+DNO/FC/HBET:3	152	4,901	$ 425	$ 2,082,802
RC4	RC walls without ERD	/CR+CIP/LWAL+DNO/FC/HBET:8	190	14,677	$ 425	$ 6,237,659
RCPC1	RC walls with moderate level of ERD	/CR+PC/LWAL+DUC/FC/HBET	152	4,910	$ 425	$ 2,086,638
RCPC2	RC walls with high level of ERD	/CR+PC/LFLS+DUC/FC/HBET:5	152	5,112	$ 425	$ 2,172,612
ADO	adobe	/MUR+ADO/LWAL+DNO/FW/HBET:1	5.2	103	$ 60	$ 6,174
WOOD1	timber structure	/W/LWAL+DUC/FW/HBET:1	5.2	1,867	$ 255	$ 476,032
WOOD2	timber structure	/W+WLI/LO+DUC/FW/HBET:1	5.2	1,812	$ 100	$ 181,189
STEEL	steel structure	/S/LFM+DNO/FME/HBET:1	3.8	387	$ 175	$ 218,125
OTH	other structure	OTH	76	2,054	$ 425	$ 359,488

Mapping housing census and establishment survey data to building classes

The residential exposure models presented in all cases, except Turkmenistan, are derived using the latest published housing census tables in each country at the smallest available administrative level, and with a modular structure that allows local, national and regional improvements or updates to be incorporated, as new data is released. This is considered to be a bottom-up approach for creating the building exposure models, as opposed to top-down approaches that use aggregate statistics such as the total population or GDP per unit area to estimate the number of buildings in that area. For Afghanistan, which has not had a nationwide census since 1979, the Afghanistan Living Conditions Survey (ALCS) 2016-17[33] is used as the primary data source. The ALCS was designed to provide representative results at the national and provincial level. For Turkmenistan, where no housing information is available at the subnational level, the top-down approach is adopted, using the gridded population dataset from WorldPop[34] as a proxy for the number of buildings.

Typically, the housing census for a given country provides the number of buildings or dwellings by type of housing unit (e.g., individual house, apartment house). Less often, information about the building wall material, age of construction, height and presence of basements is also included. The data are usually found aggregated at a given administrative division (e.g., province, municipality). The information provided by the housing census varies considerably between countries. Moreover, the distinction between rural and urban areas is often not made in all housing censuses.

For the commercial and industrial building counts, there are no national surveys that compile specific building information in any of the member countries under consideration, and therefore secondary data sources were used, such as the establishment census, statistical yearbooks and labor force (number of employees). In all cases except Turkmenistan, the existing databases included the number of establishments, size of the enterprises (or number of employees) and the principal economic activity (e.g., retail, manufacturing, warehouse, construction, mining) at the smallest available administrative level. For Turkmenistan, where this data was not available, proxies of distributed workforce across economic sectors were

[33] Central Statistics Organization (2018), Afghanistan Living Conditions Survey 2016-17. Kabul, CSO. Available at: https://www.nsia.gov.af:8080/wp-content/uploads/2020/10/ALCS-2016-17-Analysis-report-English-23-.09-2018_compressed-1.pdf
[34] WorldPop. 2021. Accessible at: https://www.worldpop.org/methods/populations

combined with national and regional estimates of average area per worker, and average area per facility per economic sectors, to generate the distribution of buildings per economic sector.

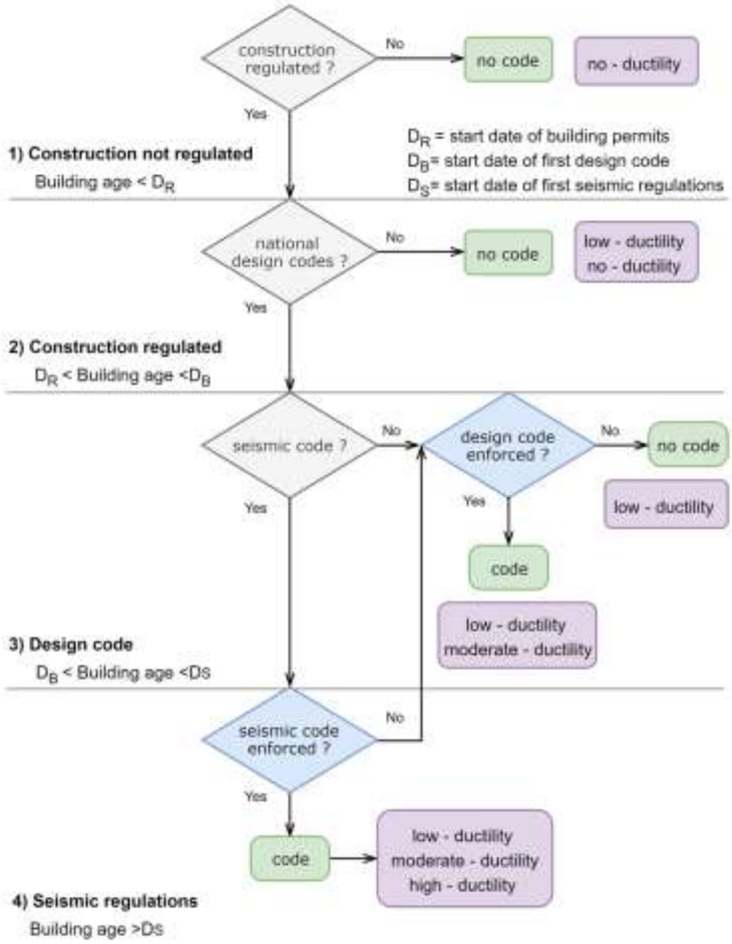
Each country presents the relevant census databases with different levels of detail, formats and spatial resolution, and census databases generally focus on population dynamics, housing conditions, economic activity and employment characteristics and indicators, rather than on the number of buildings in different vulnerability classes directly. The definition of the number of existing buildings per occupancy type plus the categorization into building classes require additional efforts that incorporate uncertainties in the modeling process.

For the Xinjiang Uyghur and Inner Mongolia Autonomous Regions of PRC, the primary reference datasets for the residential exposure are the household statistics tables published by the Sixth National Population Census of the People's Republic of China, conducted in 2010. The datasets compiled by the census using the long-form survey include information about the characteristics of households in rural, town, and urban areas. This information is aggregated up to the various administrative levels. While the province-level and prefecture-level household statistics tables are available for all provinces, county-level aggregated data are not available for a few provinces. The main household characteristics that are of interest for the development of a building exposure model —including the building bearing type, the range of the number of stories, and the year of construction— are obtained from the long-form section of the 2010 census. Additional characteristics, such as the average size of family households and the average floor space per capita, are obtained from the detailed census tables.

The building bearing type in the PRC census data is coarsely divided into “Steel and Steel Reinforced Concrete”, “Composite Structure”, “Brick-Wood Structure”, and “Other”. For the purposes of the assignment of appropriate structural vulnerability functions to the buildings, this field must be further subdivided into structural classes, such as wood-frame structures, earthen or adobe masonry, stone masonry, brick masonry (unconfined and reinforced), reinforced concrete (infill-wall moment frame, shear-wall, and dual-system) structures, and steel structures. This disaggregation step is informed by consultation with local engineering experts, the existing literature and author knowledge. The disaggregation of households by building bearing type into households by structural type is undertaken separately for the rural, town, and urban areas since construction practices and dominant technologies are often different in the three cases.

Another attribute that is often challenging to define is the ductility level of buildings, which is frequently used to describe the expected seismic performance of the building. Ductility levels are assigned based on the type of construction, age of construction and code enforcement efficiency. In this study, four levels of ductility were considered: non-ductile, low, moderate and high ductility. For the identification of the ductility, it is necessary to understand the history of seismic design implementation in each country, and this involves research about the existing seismic design codes in the different countries and the extent of their enforcement (if applicable). Figure 6 shows an overview of the workflow used for assigning seismic ductility levels to the buildings in the different countries.

Figure 6: Workflow for the assignment of ductility levels



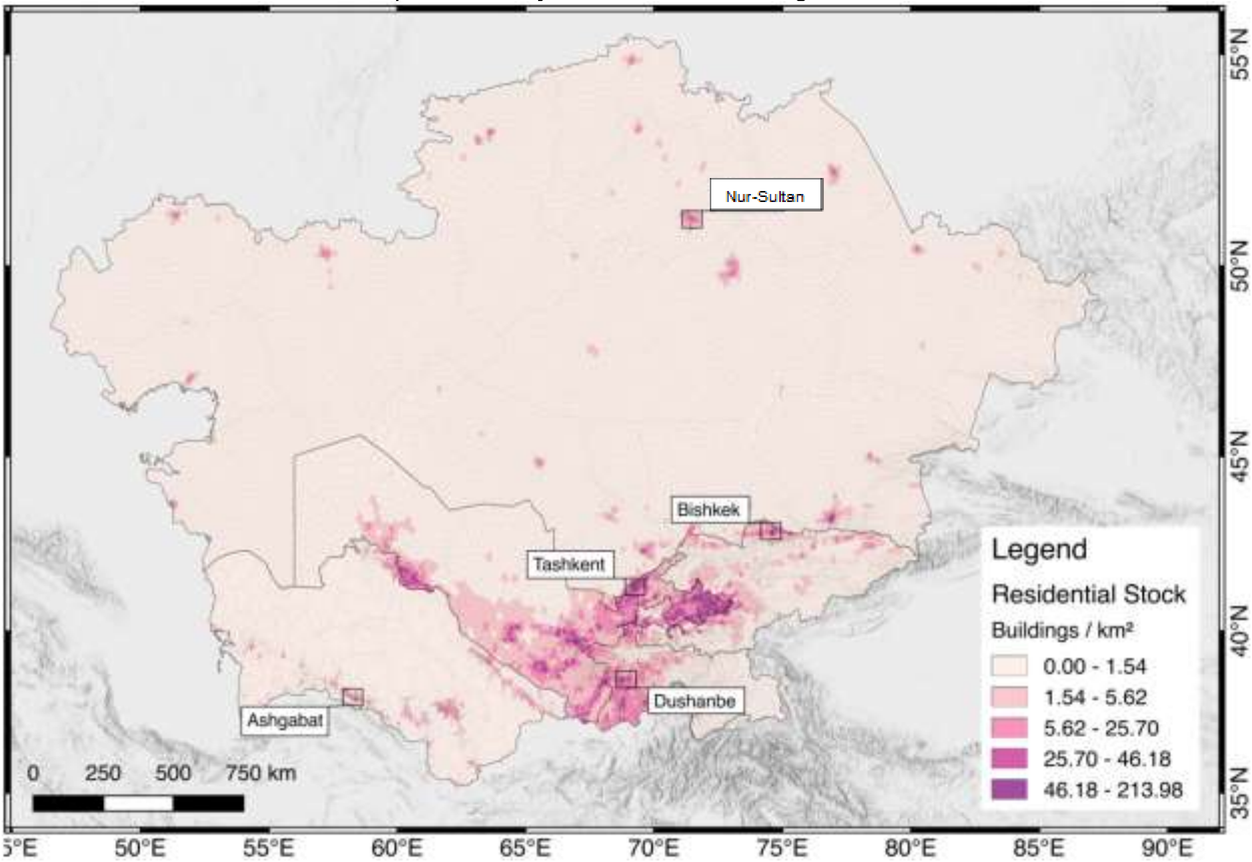
Source: Dabeek and Silva, 2019

Mapping housing units or establishments to buildings

The method used for the estimation of building counts depends on the occupancy type, since the residential and non-residential (commercial and industrial) data sources vary considerably. For the residential buildings, the primary data sources were the population and housing censuses, which are generally updated every 10 years and follow international principles and standards. For each country, geo-referenced information at the smallest administrative level was collected, combining as many relevant variables as possible, such as the main construction material of the outer walls (e.g. masonry, concrete, wood, palm leaves), the dwelling type (e.g. single family house, apartment, hut), the number of storeys, the material of the floor (e.g. tiles, earth, concrete), the material of the roof (e.g. concrete slab, wood, thatch roof), the settlement type (urban or rural areas), year of construction, and in rare cases, information regarding the state of conservation of the building (Figure 7).

Typically, information in the housing census is expressed in terms of dwellings or buildings counts. The number of dwellings needs to be converted to an estimate of the number of buildings using the average number of floors per building and the average number of dwellings per floor, for the various building types considered in the exposure model.

Figure 7: Residential building exposure model for the Central Asia region showing the estimated spatial density of residential buildings



Source: Pittore et al., 2020.

Estimation of built-up areas and replacement costs

Quantifying the floor area and replacement cost of buildings in the exposure model is essential for the calculation of economic losses. The definition of the average built-up area and average replacement cost depends on the building class, the occupancy type, and for the residential buildings, on the settlement type (urban or rural). For Mongolia and the PRC, the National Statistical Offices reported average sizes and costs as part of their national census reports, whereas for the rest of the CAREC member countries where national data was not available, regional, and global valuation surveys were used as reference values. The replacement costs reported in the model are in US dollars and have been adjusted to present values. The costs included the structural and non-structural components of the buildings and its contents and were estimated on the basis that all buildings should be replaced by formal construction in accordance with national construction standards, and not the existing conditions.

For the Central Asian countries of Kazakhstan, the Kyrgyz Republic, Tajikistan, Turkmenistan, and Uzbekistan, the average floor areas are derived using building footprints extracted from OpenStreetMap, from a sample of around 7,000 surveyed buildings in the Kyrgyz Republic and Tajikistan.[35] Estimates for the average unit replacement costs are based on the average floor area and average number of stories per building of a particular class. The OTH (other) class has been assigned by default for the replacement cost of unreinforced masonry, based

[35] Pittore, M., Haas, M., & Silva, V. (2020). Variable resolution probabilistic modeling of residential exposure and vulnerability for risk applications. Earthquake Spectra, 36(1_suppl), 321–344. https://doi.org/10.1177/8755293020951582

on the survey results. The replacement costs for each of the building classes are listed in Table 3 above.

For Azerbaijan, Georgia, Pakistan, and Afghanistan, for the purposes of assigning replacement costs based on the quality of construction, two levels of construction quality were assumed conditional on the main construction material: reinforced concrete is considered to have intermediate quality, while other materials are deemed to have a lower quality.

Replacement costs for Mongolia for different types of construction, including gers, apartments, and detached houses, are derived from price tables in the national report of the 2020 Population and Housing Census of Mongolia.[36]

The main source used for replacement costs for the Xinjiang Uyghur and Inner Mongolia Autonomous Regions is the 2010 China Statistical Yearbook[37], using the floor space of residential buildings completed and the total value of residential buildings to infer the average construction cost. An annual construction cost escalation rate of 3% was assumed to bring the value per square meter in 2010 to the present value (2020). The replacement cost was compiled separately for the urban, town, and rural areas of each of the two provinces. It represents the average value of residential construction, including both structural components and nonstructural components, but it does not include building contents. The average size per dwelling and floor space per person are used to estimate the average floor areas of all houses and buildings, and the average cost per unit area is applied to estimate the total replacement cost for the buildings. The average size of dwellings and average cost per unit area for houses in urban, semi-urban, and rural areas for the two autonomous provinces are listed in Table 5.

Table 5: Average size of dwellings and average cost per unit area for houses in urban, semi-urban, and rural areas in Xinjiang Uyghur and Inner Mongolia Autonomous Regions of the PRC

Autonomous Region	Urban				Semi-Urban				Rural			
	Value of Houses			Avg Floor Space per Capita (m²)	Value of Houses			Avg Floor Space per Capita (m²)	Value of Houses			Avg Floor Space per Capita (m²)
	(¥/m², 2010)	(¥/m², 2019)	(US$/m², 2019)		(¥/m², 2010)	(¥/m², 2019)	(US$/m², 2019)		(¥/m², 2010)	(¥/m², 2019)	(US$/m², 2019)	
Nèi Měnggǔ Zìzhìqū (Inner Mongolia)	¥ 2,700	¥ 3,163	$ 455	31.8	¥ 2,350	¥ 2,753	$ 396	31.5	¥ 600	¥ 703	$ 101	29.9
Xīnjiāng Wéiwú'ěr Zìzhìqū (Xinjiang)	¥ 1,800	¥ 2,109	$ 303	36.2	¥ 1,800	¥ 2,109	$ 303	33.4	¥ 600	¥ 703	$ 101	29.6

Uncertainties and limitations in the data and methodology

Overall, the exposure data includes residential, commercial, and industrial built assets for all member countries, with two exceptions: Mongolia and Turkmenistan have a residential exposure model alone, as insufficient information was available for modeling the commercial and industrial exposure. Building and population counts in the exposure models are based on the last available census for each country, except for Afghanistan, Mongolia, and Turkmenistan, as explained earlier which do not have a recent census.

Uncertainties remain within the exposure model development, particularly in identifying the location of the assets, definition of the structural attributes of the assets including the building typologies and approximate period of construction, estimating the floor-areas and replacement costs of buildings.

Exposure data is currently unavailable for certain regions of Azerbaijan, Georgia, and Pakistan. For example, data for the Kalbajar-Lachin region is not available in the Azerbaijan census. In the case of Georgia, the 2014 census could not be carried out in the territories of Abkhazia and South Ossetia. Additionally, the northern parts of the region of Shida Kartli are

[36] National Statistics Office (NSO) of Mongolia. 2020 Population and Housing Census of Mongolia. National Report.
[37] China Statistical Yearbook. Accessible at: http://www.stats.gov.cn/tjsj/ndsj/2010/indexch.htm

absent from the national census data. Finally, the exposure model for Pakistan is based on the 2017 housing census. At the time of modeling, the housing census datasets for the territories of Gilgit-Baltistan and Azad Jammu and Kashmir had not been released yet. A census has been conducted for these territories, but the datasets are still not published. For the above-mentioned territories/regions of Azerbaijan, Georgia, and Pakistan, exposure is modelled using population estimates for the respective regions from WorldPop as a proxy to estimate the number of buildings.

Vulnerability

Data sources and methodology

Earthquake

Economic damage

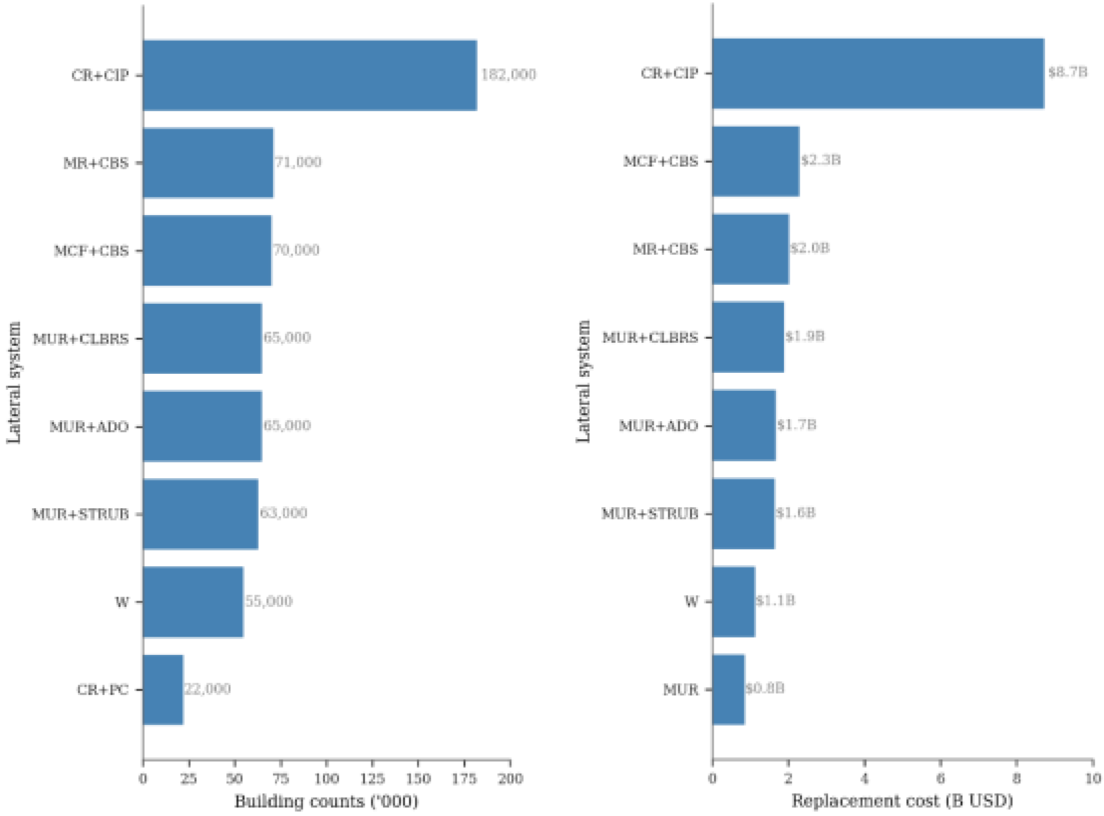
The vulnerability of assets to earthquake hazards is tightly linked to the characteristics of the assets themselves. Accordingly, the GEM Global Asset Database building classifications described in the exposure section consider the main physical characteristics that influence the expected performance of buildings under seismic loads.

For the Central Asian countries of Kazakhstan, the Kyrgyz Republic, Tajikistan, Turkmenistan, and Uzbekistan, the exposure model defines fifteen building vulnerability classes, as shown in Table 3.[38] This classification was informed by a set of remote mapping surveys conducted by local engineers in the Kyrgyz Republic and Tajikistan between 2012 and 2016.[39,40] The residential building stock in these countries is dominated by rural construction, mostly using adobe, unreinforced masonry, and wood as the preferred material of construction. Buildings that could not be clearly assigned to any one typology due to ambiguity have been assigned to a sixteenth ‘OTH’ (other) class.

For Azerbaijan, Georgia, Pakistan, and Afghanistan, the predominant construction classes were identified from a review of existing literature, studies describing the architecture and energy consumption, and solicited feedback from local engineers. Thirty unique building classes were determined according to the material of construction, lateral load resisting system, ductility level, and number of stories. An example of how vulnerability varies by occupancy type is provided in Figure 8 below.

[38] Pittore, M. Haas, M., Silva, V. (2020) Variable resolution probabilistic modeling of residential exposure and vulnerability for risk applications, *Earthquake Spectra,* 36(1), 321-344.
[39] Megalooikonomou, KG, Parolai, S, Pittore, M (2018) Toward performance-driven seismic risk monitoring for geothermal platforms: Development of ad hoc fragility curves. *Geothermal Energy,* 6: 8.
[40] Pittore, M, Haas, M, Megalooikonomou, KG (2018) Risk-oriented, bottom-up modeling of building portfolios with faceted taxonomies. *Frontiers in Built Environment,* 4: 41.

Figure 8: Occupancy types based on physical vulnerability to earthquakes for Georgia.



Source: Global Earthquake Model. Explanation of the GEM Taxonomy strings: CR+CIP = Cast-in-place reinforced concrete; MR+CBS = Reinforced masonry using solid concrete blocks; MCF+CBS = Confined masonry using solid concrete blocks; MUR+CLBRS = Unreinforced masonry using fired clay bricks; MUR+ADO = Unreinforced adobe masonry; MUR+STRUB = Unreinforced masonry using rubble or semi-dressed stones; W = Wood frame; CR+PC = Precast concrete; MUR = Unreinforced masonry (general)

People affected and fatalities

The number of people affected by earthquakes is defined as the population that can be expected to witness earthquake-caused ground shaking of Modified Mercalli Intensity (MMI) VI or higher (corresponding to strong shaking, capable of causing slight damage or higher). The average annual people *severely* affected by earthquakes is defined as the population that can be expected to witness earthquake-caused ground shaking of MMI VIII or higher (corresponding to severe ground shaking, capable of causing considerable damage including partial collapses in ordinary structures, along with slight damage to well-engineered structures).

The number of fatalities due to earthquakes is estimated using GEM's global database of analytical seismic fragility curves and vulnerability functions. Fragility curves describe the probability of exceeding a set of damage states conditional on a ground shaking intensity level, and are fundamental for the assessment of damage in earthquake scenarios. These functions are converted into fatality vulnerability functions for building occupants using a damage-to-loss model, i.e. a relation between a given damage state for a particular building class and the corresponding fatality rate[41] for the building occupants, leading to a distribution of the probability of fatality conditional on a set of ground shaking intensities.

[41] So, E. (2016). Estimating fatality rates for earthquake loss models. Springer International Publishing. 71pp.

Indirect economic damage

Indirect economic damage is defined as the additional impacts beyond direct damage that result from earthquake shaking. These relate to economic disruption from damage to infrastructure and industrial facilities and the disruption to economic operations that result (e.g., difficulties for workers accessing the workplace leading to lowered productivity). The impacts increase for more severe events due to the length of reconstruction and demands on skilled trades and materials required for reconstruction. Following Hallegatte,[42] the net present value of discounted indirect damages for earthquake was estimated using the following equation:

$(rN / rN + 3) \times (1/\mu) \times (\Delta K)$

where,

- ΔK is the direct damage or loss of capital
- r is the marginal productivity of capital
- (r/μ) is the average productivity of capital
- N is the number of years required for repair / reconstruction of >95% of the damaged assets.

The marginal productivity of capital at the national level is estimated using the most recently available real rate of interest in the country. Average productivity of capital is calculated as the country's output-side GDP divided by total reproducible capital within the country, following Hallegatte and Vogt-Schilb.[43] These two variables are both available from the Penn World Tables for all CAREC countries except Afghanistan. For Afghanistan, we assume the default value of 0.3 for the average productivity of capital, as mentioned in Hallegatte (2015). The reconstruction period as a function of the return period is estimated based on another working paper by Rozenberg and Hallegatte (2016)[44]. A high scenario is used for the CAREC countries as reconstruction is generally expected to take longer in developing countries (Table 6).

Table 6: Reconstruction time after a disaster event by return period and for the three scenarios.

	Return period			
Scenarios	0-10	10-100	100-500	>500
Low	0.5	1	2	3
Medium	1	2	3	5
High	3	3	5	10

Since several return periods are being reported in the loss exceedance curves presented in the risk profiles and modeling, a function is used to estimate reconstruction times based on the return period that yields values along the high scenario estimates from the above table:

$N = [\log_{10}(RP)]^2$

Finally, the indirect AAL is calculated as the area under the loss exceedance curve.

[42] Hallegatte, S. (2015) The Indirect Cost of Natural Disasters and an Economic Definition of Macroeconomic Resilience. *World Bank Policy Research Working Paper* https://doi.org/10.1596/1813-9450-7357
[43] Hallegatte, S. and Vogt-Schilb, A. (2016) Are Losses from Natural Disasters More Than Just Asset Losses? The Role of Capital Aggregation, Sector Interactions, and Investment Behaviors. *World Bank Policy Research Working Paper.* https://doi.org/10.1596/1813-9450-7885
[44] Rozenberg, J. and Hallegatte, S. (2016) Modeling the Impacts of Climate Change on Future Vietnamese Households: A Micro-Simulation Approach. *World Bank Policy Research Working Paper.*

Flood

Economic damage

The Global Flood Model has been developed with a set of default regional and global vulnerability functions for different lines of business, including residential, commercial and industrial exposure. These vulnerability functions describe the relationship between flood depth and damage based on empirical studies, engineering data and other sources of information. The Global Flood Model vulnerability functions were developed based on two main sources of reference:

- JBA's existing vulnerability functions developed from projects in Australia, Canada, PRC, Sierra Leone, Thailand, UK, US, and Vietnam.
- Information from a recent research report on global flood vulnerability functions published by the European Commission's Joint Research Centre (Huizinga, Moel and Szewczyk, 2017).

Initially, JBA default vulnerability functions for the region were used to generate the model outputs. The modelled results produced using the default functions were then compared with loss estimates for historic events. No vulnerability functions were found in the literature for any of the individual CAREC countries or region. As a result, modelled losses were validated against historic losses to assess the efficacy of the vulnerability relationships. Based on this comparison of modelled and recorded flood losses, and using the information available in the exposure data, the vulnerability functions were adjusted to take into account some of the specific building types in the region. Vulnerability functions were mapped to the building categories in the exposure data before model loss estimates were recalculated. Assessment of economic damage from flooding use vulnerability functions derived as part of the project. These functions determine the relationship between hazard intensity (flood depth) and damage.

People affected and fatalities

To determine the number of people affected by flooding, a threshold of 0.2m is used to set a level above which a property is considered to be flooded and the associated population (people per building) is considered to be 'affected'. Population estimates were based on a point-based distribution of residential population at 1 km resolution from Global Human Settlement Layer (GHSL) data (value in 2015). Spatial distributions at sub-national level were derived from the GHSL data for all CAREC countries. A ratio between the GHSL and 2020 UN population data was calculated at country level to scale up the spatial distribution to the 2020 population counts. These counts were then summed for each top-level administrative division polygon.

The categorization of people severely affected by flooding is used to indicate where flood impacts are sufficiently severe that properties are uninhabitable for a significant period of time (weeks or months). With this level of disruption, it is assumed that there are associated impacts on daily life (e.g., disruption to working life, temporary relocation to a different area increased travel time to work) and wellbeing (e.g., mental health impacts of dislocated communities, damage or loss of possessions). To indicate this level of impact, a second depth threshold was used in the flooding modeling, set at 1.0m, to identify those areas and properties where impacts on the population were likely to be severe.

Jonkman et al. (2008)[45] review a range of studies linking flood characteristics with fatalities. Some of these consider the many factors that relate flood hazard to fatalities, including the type of flooding, the vulnerability of different population groups, the flood mitigations in place (e.g. flood defences, warning systems) and the timescale of the event. Many of these studies relate to a specific country, making application to the CAREC region questionable. No studies were found relating flood hazard to fatalities in the CAREC region. Due to lack of data to build a complex function relating flood hazard to fatalities, an empirical approach was taken.

Fatalities from flooding have been calculated using functions that relate the cumulative number of people affected to the corresponding death toll, based on evidence from historical events for each CAREC country collected as part of the validation exercise for the flood modeling, where available. For two countries, Kazakhstan and Turkmenistan, this data was not available. In this case, the average ratio of the other nine CAREC countries was used (Table 7).

Table 7: Table of fatality ratio to people affected for flood risk across the CAREC member countries

Country	**Ratio of fatalities to people affected for flood**
Afghanistan	0.0005
Azerbaijan	0.0001
People's Republic of China	0.00005
Georgia	0.0068
Kazakhstan	0.0025
Kyrgyz Republic	0.0071
Mongolia	0.0053
Pakistan	0.0001
Tajikistan	0.0015
Turkmenistan	0.0025
Uzbekistan	0.001

Indirect economic damage

Indirect economic damage is defined as the additional impacts beyond direct damage that result from flooding. These relate to economic disruption from damage to infrastructure and industrial facilities and the disruption to economic operations that result (e.g., difficulties for workers accessing the workplace leading to lowered productivity). The impacts increase for more severe events due to the length of reconstruction and demands on skilled trades and materials required for reconstruction. Following Hallegatte,[46] the net present value of discounted indirect damages for flooding was estimated using the following equation:

$(rN / rN + 3) \times (1/\mu) \times (\Delta K)$

where,

- ΔK is the direct damage or loss of capital
- r is the marginal productivity of capital
- (r/μ) is the average productivity of capital
- N is the number of years required for repair / reconstruction of >95% of the damaged assets.

[45] Jonkman, S.N., Vrijling, J.K. and Vrouwenvelder, A.C.W.M. (2008) Methods for the estimation of loss of life due to floods: a literature review and a proposal for a new method. *Natural Hazards*, vol. 46, pp.353-389
[46] Hallegatte, S. (2015) The Indirect Cost of Natural Disasters and an Economic Definition of Macroeconomic Resilience. *World Bank Policy Research Working Paper* https://doi.org/10.1596/1813-9450-7357

The marginal productivity of capital at the national level is estimated using the most recently available real rate of interest in the country. Average productivity of capital is calculated as the country's output-side GDP divided by total reproducible capital within the country, following Hallegatte and Vogt-Schilb.[47] These two variables are both available from the Penn World Tables for all CAREC countries except Afghanistan. For Afghanistan, we assume the default value of 0.3 for the average productivity of capital, as mentioned in Hallegatte (2015). The reconstruction period as a function of the return period is estimated based on another working paper by Rozenberg & Hallegatte (2016)[48]. A high scenario is used for the CAREC countries as reconstruction is generally expected to take longer in developing countries (Table 8).

Table 8: Reconstruction time after disaster by return period and for the three scenarios.

	Return period			
Scenarios	0-10	10-100	100-500	>500
Low	0.5	1	2	3
Medium	1	2	3	5
High	3	3	5	10

Since several return periods are being reported in the loss exceedance curves presented in the risk profiles and modeling, a function is used to estimate reconstruction times based on the return period that yields values along the high scenario estimates from the above table:

$N = [\log_{10}(RP)]^2$

Initially, default values proposed in the Hallegatte paper were used for the function. Because of lack of information on the disproportionate impacts to critical infrastructure, the “ripple effect” was set to zero. Interest rates for the CAREC countries lie in the range 4 to 16%, but it was decided to set this value at 20% to indicate the overall potential lost capital opportunities and the imperfect operation of capital markets.

One of the most uncertain elements of the calculation is the reconstruction period. Three years is proposed in Hallegatte (2015), but, in other literature, recovery periods between 6 months and several years are suggested for a range of disaster scenarios such as the 2004 Asian tsunami. Selecting three years for the reconstruction period resulted in very high values of indirect losses for higher return period events, when compared to analyses for specific case studies such as the Mumbai floods of 2005 (return period estimated at 100 years) and hurricane Katrina in the same year (RP ~25 years). To account for the change in reconstruction and recovery period for events with higher return periods, and the range of estimates from the literature, an adjusted function was developed for calculating indirect economic loss. Finally, the indirect AAL is calculated as the area under the loss exceedance curve.

Infectious Disease

The application of catastrophe modeling techniques to the analysis of infectious disease requires the development of complex, dynamic models, which differ substantially from traditionally natural catastrophe models (such as earthquake and flood) in terms of how elements such as hazard and vulnerability are conceptualized, measured, and modeled.

Vulnerability within the infectious disease modeling framework is related to a number of factors, including the attributes of the specific pathogen being modeled, and country-specific characteristics such as preparedness and policy responses. Within the Metabiota disease

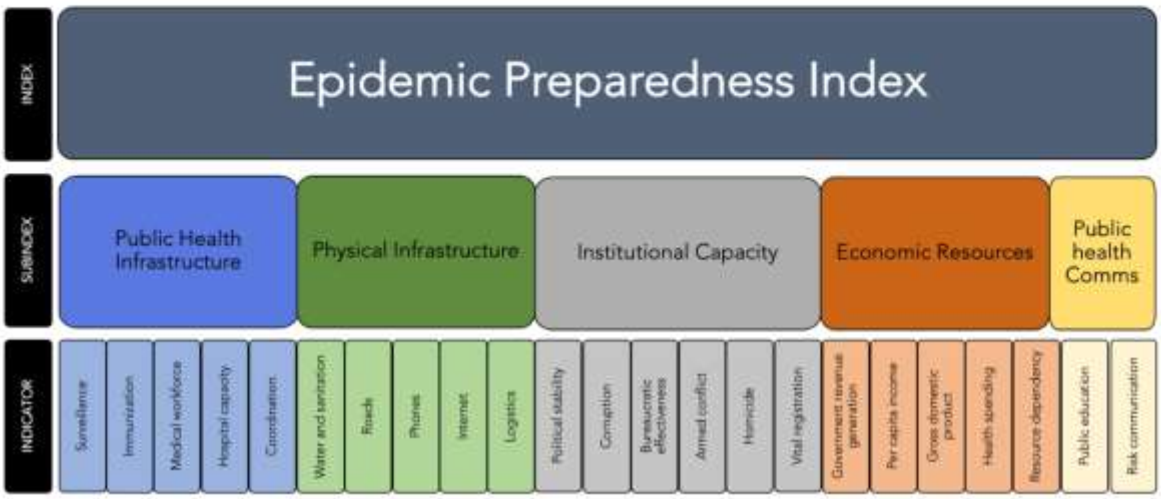
[47] Hallegatte, S. and Vogt-Schilb, A. (2016) Are Losses from Natural Disasters More Than Just Asset Losses? The Role of Capital Aggregation, Sector Interactions, and Investment Behaviors. *World Bank Policy Research Working Paper.* https://doi.org/10.1596/1813-9450-7885

[48] Rozenberg, J. and Hallegatte, S. (2016) Modeling the Impacts of Climate Change on Future Vietnamese Households: A Micro-Simulation Approach. *World Bank Policy Research Working Paper.*

spread model, vulnerability depends on pathogen-level characteristics including its case fatality ratio (overall, and for specific age bands), transmission dynamics, and the availability of vaccines or therapeutics. Vulnerability is also dynamic within each simulated event, changing on a daily time-step within the model: for example, individual vulnerability may decrease due to immunization or infection that confers immunity within the simulation, while localized population vulnerability may decline as the number of susceptible individuals within a simulated geographic area declines.

Vulnerability at a country level is also influenced by preparedness to detect and respond to health emergencies: to identify a potential infectious disease risk, implement appropriate policy responses (such as social distancing), and, if feasible, implement pharmaceutical interventions. As noted above, preparedness is measured via an epidemic preparedness which takes into account both health system capacity as well as factors which enable (or constrain) effective response, including governance, financing, risk communications, and physical and communications infrastructure.

Figure 9 Epidemic Preparedness Index Design[49]



Within the modeling framework, each country's epidemic preparedness index score influences the speed with which it implements non-pharmaceutical responses, and (if vaccines are relevant for the pathogen being modeled) the timing and rate at which it begins vaccination efforts. This creates substantially different vulnerability profiles across countries and over time.

Uncertainties and limitations in the data and methodology

The sparsity of building vulnerability surveys across the region contributes to uncertainty in earthquake vulnerability data. For instance, surveys undertaken in the Kyrgyz Republic and Tajikistan were used to inform the curves used here and in the surrounding countries of Kazakhstan, Turkmenistan, and Uzbekistan. Furthermore, in some cases building characteristics did not map neatly onto vulnerability typologies necessitating the use of an 'other' vulnerability class. Where field surveys could not be consulted directly, relevant literature was sought and applied.

[49] Oppenheim, B., Gallivan, M., Madhav, N. K., Brown, N., Serhiyenko, V., Wolfe, N. D., & Ayscue, P. (2019). Assessing global preparedness for the next pandemic: development and application of an epidemic preparedness index. *BMJ global health*, *4*(1), e001157.

No flood vulnerability functions were found in the literature covering the CAREC region. Accordingly, JBA default vulnerability functions were used to generate initial modelled losses and then calibrated against loss estimates for historic events. Uncertainties may arise during this approach given that there are relatively few historical loss events that are suitable for calibrating vulnerability curves. That said, the approach used was judged to be the most appropriate way to develop tailored vulnerability curves across the CAREC region given limited data availability.

When calculating indirect economic damages from both earthquake and flooding, uncertainty arises from the use of the use of macroeconomic indicators (real rate of interest, GDP, total reproducible capital) which in reality would vary spatially across the country / region. Default values for these macroeconomic indicators were used in some cases (e.g., Afghanistan). Similarly, assumptions were necessary when quantifying the 'ripple effect', interest rates, and reconstruction periods. Reconstruction periods in particular can vary from months to years and have a marked impact on indirect damages.

Assessing the protection gap across the CAREC region

Defining the protection gap

The protection gap is traditionally defined as the proportion of losses from disaster events that are not insured. Identifying the level of risk which has not been reduced (through risk reduction investment) or transferred (through risk financing) is to identify the contingent liability that will need to be met in the event of a disaster. This is used as a fundamental input into the design of risk management and arrangement of risk financing.

Data sources and methodology

An assessment and quantification of the protection gap was undertaken across the CAREC region. The risks associated with earthquake and flood in each of the CAREC countries / regions was compared with the financing approaches available to governments, households, and businesses to help identify where and what type of additional disaster risk financing may be most valuable at the country and regional scale. The outcomes from this analysis are presented in the Protection Gap Assessment report. This section details the methodology that underpins that report (Figure 10).

Figure 10 Assessment and quantification of the protection gap across the CAREC region.

i. apply catastrophe risk modeling to understand the direct and indirect economic losses for each country

↓

ii. review existing ex-ante risk financing arrangements

↓

iii. compare the gap between the modeled losses and ex-ante arrangements

↓

iv. consider other available sources of financing to close the gap

↓

v. summarise key insights for each country and the implications for involvement in a regional risk transfer facility

The probabilistic risk modeling completed under this Technical Assistance (and described above) was used to assess the direct losses and indirect costs face by each country / region. In this case, direct losses include damage to commercial, industrial, and residential property, while indirect costs consider ongoing reductions in business / household income due to disruptions (e.g., in supply chains). The human costs of earthquake and flooding events are also considered.

Alongside country / regional level analysis, a brief sub-national analysis was also completed, comparing estimates of expected economic and human cost with sub-national data on poverty[50], recognizing that poverty is likely to accentuate the vulnerability of the population to

[50] Where data is available, the sub-national estimates of poverty are taken from the Multidimensional Poverty Index from the Oxford Poverty and Human Development Initiative (https://ophi.org.uk/multidimensional-poverty-

disaster events in ways that are not fully captured by catastrophe modeling. This allows the initial identification of sub-national 'hotspots' where both hazard risk is high and poverty, relative to elsewhere in the country, is also high.

A qualitative assessment of the ex-ante risk reduction and risk financing mechanisms in place was performed. This includes looking both at provisions made at the sovereign and sub-sovereign level, and the extent to which insurance solutions are adopted across different groups in society.

The 'protection gap' is quantified as the difference between (i) the losses / costs faced by a given country, and (ii) the financing mechanism available to cover them. Two comparisons are made:

- To determine the difference between average annual losses and the cover offered by existing risk retention and transfer mechanisms; and
- To establish the return period loss events that, when accounting for losses that might be insured, would exhaust the available risk retention mechanisms

In the second case, the 'losses that might be insured' are examined in three ways:

- assuming that the risk retention mechanisms need to cover both the uninsured direct losses and indirect costs (as described above);
- assuming that risk retention mechanisms need to cover only the uninsured direct losses that events might cause; and
- assuming that risk retention mechanisms need to cover the emergency response costs associated with different events.[51]

Other sources of finance available to governments, households, and businesses to meet the financial costs of disaster events were then considered. At the sovereign level, this involves an examination of the macroeconomic context of the country and, where data is available, its financial response to previous disaster events. At the level of households and businesses, the assessment is predominantly carried out through a review of the extent of financial inclusion in the country which provides a proxy for their ease of accessing finance. This is complemented by an assessment of social protection provision, recognizing that social protection is often an essential way for vulnerable individuals and households to cope with disaster events[52], while recognizing at the same time that such mechanisms will also increase contingent liabilities at the sovereign level.

Finally, each of the above steps are drawn together to provide a summary of the key insights for each country and the implications for involvement in a regional risk transfer facility.

index/#:~:text=The%20global%20Multidimensional%20Poverty%20Index,that%20a%20person%20faces%20simultaneously.) This is the preferred data sources as it complements monetary poverty measures by capturing deprivations in health, education and living standards that a person faces simultaneously. The dataset also provides information on the number/proportion of people living in poverty. However, this dataset is not available for all CAREC countries so, where necessary, data on the human development index score at the sub-national level was used instead (https://globaldatalab.org/shdi/shdi/). The human development index aggregates information on life expectancy, education and per capita income and converts this to an index score of between 0-1 with scores closer to 1 representing higher levels of human development.

[51] It is assumed that emergency response costs are equal to 23% of the direct losses of a flood event and 16% of the direct losses of an earthquake event, based on industry practice derived from an analysis of historic events. Insurance penetration is not taken into account when assessing the extent to which risk retention will cover emergency response costs.

[52] World Bank (2019) Social Protection and Disaster Recovery. Accessible at: https://www.gfdrr.org/sites/default/files/publication/Social_Protection_Guidance_Note_FINAL.pdf

Modelling disaster risk reduction and climate adaptation measures

As part of this TA, disaster risk reduction and climate adaptation modeling will be performed using the CLIMate ADAptation (CLIMADA) modeling platform.[53] In addition to facilitating dialogue with the CAREC member states throughout the duration of this project and beyond, this work will demonstrate more generally the value of CLIMADA in taking the outputs of fully probabilistic risk models and investigating the comparative benefit of alternative disaster risk reduction and climate adaptation measures.

Data sources and methodology

CLIMADA provides a framework for users to combine exposure, hazard and vulnerability data to calculate risk. It is an open-access model that is designed to be highly flexible to user need. Furthermore, CLIMADA includes the ability to conduct cost-benefit analysis of varied disaster risk reduction and adaptation measures. A CLIMADA model has been developed to enable country-specific assessment of earthquake and flood risk and to compute the relative cost-benefit of a selection of disaster risk reduction measures.

The CLIMADA model developed for this project combines earthquake and flood hazard maps generated by the JBA and GEM fully probabilistic source models with point-based build asset data, and country-specific vulnerability curves.

Table 9 shows the areas of concentrated exposure that are included within the CLIMADA model. Most locations are country capitals, to capture areas with the highest value at risk. In some countries alternative urban areas have been selected due to nature of the hazard (e.g., Pakistan) or exposure (e.g., Kazakhstan).

Table 9 Areas of concentrated exposure included within the CLIMADA modeling framework.

Country	**Area of concentrated exposure**	
	Flood	**Earthquake**
Afghanistan	Kabul	Kabul
Azerbaijan	Baku	Baku
Georgia	Tbilisi	Tbilisi
Kazakhstan	Almaty	Almaty
Kyrgyz Republic	Bishkek	Bishkek
Mongolia	Ulaanbaatar	Ulaanbaatar
Pakistan	Karachi	Islamabad
PRC, Inner Mongolia Autonomous Region	Baotou	Baotou

[53] Aznar-Siguan, G. and Bresch, D.N., 2019. CLIMADA v1: a global weather and climate risk assessment platform. Geosci Model Dev.12(7), 3085-3097. https://doi.org/10.5194/gmd-12-3085-2019

PRC, Xinjiang Uyghur Autonomous Region	Ürümqi	Ürümqi
Tajikistan	Dushanbe	Dushanbe
Turkmenistan	Ashgabat	Ashgabat
Uzbekistan	Tashkent	Tashkent

CLIMADA collates various input datasets which are stored within the CLIMADA model structure. All datasets except the flood and earthquake hazard layers are contained within an Excel workbook (.xlsx) and can be readily modified. The simplicity of this set-up ensures continued usability of the CLIMADA model beyond the completion of this TA. Table 10 summarizes the inputs to the CLIMADA model.

Table 10 Summary of datasets to be integrated using the CLIMADA platform.

Type	**Dataset(s)**	**Source**	**Format**
Hazard	Flood hazard layers at six return periods: 20-yr, 50-yr, 100-yr, 200-yr, 500-yr, 1500-yr.	JBA	Raster (GeoTIFF, a spatial dataset containing hazard values at a constant resolution)
	Earthquake hazard layers at six return periods: 50-yr, 100-yr, 200-yr, 500-yr, 1000-yr, 2500-yr.	GEM	
Exposure	Building structural asset values.	Oasis Exposure Data standard format.	Comma separated values (CSV)
Vulnerability	Damage function comprising intensity (flood depth, m) and damage ratio.	JBA	Comma separated values (CSV)
	Damage function comprising intensity (peak ground acceleration, g) and damage ratio.	GEM	
Adaptation Measures	Selection of adaptation measures and details of their impact upon the hazard, exposure, and/or vulnerability.	UNU	Excel workbook (.xlsx)
Climate scenarios	Future climate scenarios to 2050 for Representative Concentration Scenario (RCP) 4.5 and 8.5.	ODI	Excel workbook (.xlsx)
Socio-economic scenarios	Future socio-economic scenarios based on GDP and or population change.	World Bank	Excel workbook (.xlsx)

The hazard inputs to CLIMADA are return period hazard maps for flood and earthquake, generated using JBA and GEM fully probabilistic source models. Rather than representing specific events, these return period maps show the probability of a given pixel experiencing, for example, a 1 in 200-year event. The return period maps are provided for flood and earthquake at six return periods.

The exposure or 'asset' input to CLIMADA comprises a point-based dataset associated with commercial, residential, and industrial building values. The exposure dataset is based on the GEM Global Exposure Database. The exposure database contains information regarding the number of buildings, geographical location, replacement costs (including the structural and non-structural components, and the building contents), number of occupants and vulnerability classes of the building stock. The GEM Building Taxonomy (Version 2.0) was used to classify the building stock in the CAREC member countries. Each exposure point has a code which allows it to be associated with an appropriate impact function for flood and earthquake hazard.

Vulnerability or 'damage function' information was collected for each area of concentrated exposure. A single damage function was used for flood risk modeling, based on the vulnerability of residential buildings in each country. For the earthquake risk modeling, appropriate damage functions were selected based on the occupancy class (commercial, industrial, residential) and building structural characteristics (e.g. reinforced concrete, unreinforced masonry, etc.).

Climate scenarios can be applied to the flood risk analysis based on observed and projected rainfall intensity duration frequency curves. Regional Climate Model-Global Climate Model (RCM-GCM) simulations from the Coordinated Regional Climate Downscaling Experiment (CORDEX) were used to examine climate change impacts on precipitation. This included two RCM-GCM simulations from the CORDEX Central Asian domain, six models from the CORDEX South Asia domain, and four models from the CORDEX East Asia domain. Two Representative Concentration Pathways (RCP4.5 and RCP8.5) were selected; these respectively represent a medium and high (business-as-usual) emissions pathway. The RCMs were bias corrected before precipitation projection analysis of how conditions could shift between the 2050s (2031-2070) and a historical reference period of 1956-1995.[54] The multi-model mean information was used to examine yearly and seasonal changes under RCP4.5 and RCP8.5. No future climate change can be applied to earthquake risk modeling.

Future socio-economic scenarios can be applied to both the flood and earthquake risk modeling. Future socio-economic change is used to alter the future exposure values based on GDP growth rate and population change projections. Future economic growth rate was estimated by calculating the mean GDP growth rate (constant to 2010 prices) over the period 2000-2019.[55] Future population growth rates are from the Population Division of the Department of Economic and Social Affairs of the United Nations.[56]

[54] The historical reference period of 1956-1995 was used over the standard 30-yr period 1961-1990 because climate over Central Asia is modulated by the Atlantic Multidecadal Oscillation and this reference period is long enough to cover two phases of the AMO, among other multidecadal climate processes. The 2050s (period 2031-2070) were chosen for the flood model (and climate modeling) as a more policy relevant period than the more distant 2070s, and a climate change signal is detectable.
[55] World Bank. 2021. GDP (constant 2010 US$). Accessed at: https://data.worldbank.org/indicator/NY.GDP.MKTP.KD
[56] Population Division of the Department of Economic and Social Affairs of the United Nations. 2018. World Population Prospectus 2018. Accessed at: https://population.un.org/wup/

Adaptation measures were selected from a set of representative adaptation options compiled by the CLIMADA development team at the United Nations University (UNU) (Table 11). These adaptation options can be readily modified to reflect particular plans in any given area of concentrated exposure.

Table 11 Description of earthquake and flood disaster risk management / climate adaptation measures as modelled using CLIMADA.

	Earthquake Adaptation Measures	**Flood Adaptation Measures**
1	*Building codes.* Assuming that new buildings comply with a given level of earthquake resistant building code.	*Channel maintenance.* Includes removing sedimentation / deepening and widening channel / clearing debris.
2	*Retrofitting.* Applying measures that elevate the earthquake resistance of existing buildings.	*Ecological restoration.* For example, increasing permeable surface cover, planting trees and restoring waterways.
3	*Insurance.* Applied as a simple sovereign insurance cover that attaches at approx. 1 in 100-year loss and exhausts at approx. 1 in 200-year loss.	*Flood awareness.* For example, campaigns to educate people about actions that can be undertaken to reduce flood impacts.
4		*Waste management.* Implementing approaches to deal with wastewater flows.
5		*Insurance.* Applied as a simple sovereign insurance cover that attaches at approx. 1 in 25-year loss and exhausts at approx. 1 in 100-year loss.

Benefit-cost ratios for each of the DRR measures are computed as total averted loss divided by total measure cost over a defined time period, in this case from present to 2050. Benefits accrue over the period from present to 2050, taking into account a discounting rate (set at default of 2%), while costs are assumed at net present value. The baseline case assumes constant exposure and hazard into the future.

The modeling undertaken here quantifies both earthquake and flood risk across the CAREC countries and presents the relative benefit-cost of a selection of representative adaptation measures. Such information helps to demonstrate the value of further detailed disaster risk reduction modeling and can contribute towards a broad-based disaster risk management agenda.

The CLIMADA model developed during this TA will remain functional beyond project termination. The legacy CLIMADA model will retain full flexibility, enabling the end-user to upload revised hazard, exposure and vulnerability data, to alter model assumptions (around future exposure growth, for example), and to amend or add to the current disaster risk reduction and adaptation measures. A separate technical guide has been prepared to accompany the legacy CLIMADA model. Full model documentation is available at: https://climada-python.readthedocs.io/en/stable/index.html.

Uncertainties and limitations in the data and methodology

The disaster risk reduction and climate adaptation modeling undertaken for this TA covers twelve cities across the eleven CAREC member countries and regions. Given this extensive geographic coverage, the modeling undertaken in any one given city is necessarily illustrative. When interpreting model outputs, it is therefore important to be aware of certain uncertainties and limitations:

- This modeling covers selected areas of concentrated exposure. Model outputs provide an indication of the risk reduction value of selected measures over a limited area and should not be used to make statements at the country or regional scale.
- Hazard input is based on a series of return period hazard maps. CLIMADA then interpolates between the input return periods to calculate intermediate losses. While this is a reasonable approach, it is not necessarily a true reflection of how losses scale with hazard intensity.
- A selection of ‘standard’ adaptation earthquake and flood measures have been modelled. The selected measures do not reflect the economic, social, political, or physical appropriateness of measures in any given city. The costs and effectiveness of adaptation measures are predicated on previous field studies conducted by the CLIMADA modeling team and adjusted for the CAREC region. The CLIMADA model does not account for the spatial distribution of adaptation measures. Further refinement over the choice, cost, and effectiveness of measures should be considered on a country basis.
- Future socioeconomic growth is based on macroeconomic indicators calculated at the country level, since 2000. In reality, past socioeconomic growth may not accurately reflect future growth and future growth may vary spatially within a country or region.

The disaster risk reduction and climate adaptation modeling undertaken here provides a consistent view of the potential effectiveness of selected measures for reducing earthquake and flood risk across the CAREC region.

Conclusion and Implications

The risk modeling undertaken as part of this TA represents the most sophisticated modeling that has been undertaken in the region to date. As a result, the regionally consistent modeling approach is valuable precisely because it appraises flood, earthquake, and infectious disease risk in each CAREC member state on the same terms. As the first attempt to quantify disaster risk in the region, it is important that this approach is open for investigation and replicable for future risk assessments. For example, the open nature of the modeling means the work conducted under the World Bank Multi-Peril Risk Assessment project, of which five are CAREC member states, is complementary. The granular nature of the World Bank project will permit review and comparison of the two approaches.

The core objectives of the risk profiling and the detail of the modeling approach in this document are (i) to improve understanding of disaster risk and disaster risk management, and (ii) to support the development of a regional disaster risk transfer facility.

One of the key challenges involved in extending high-resolution probabilistic modelling across the CAREC region is the limited availability of high-quality hazard, exposure, vulnerability, and loss data. This information is a critical input to models, as well as to guide their calibration and validation to ensure that risk is adequately captured. Improved data collection, storage, and sharing procedures should be a priority going forward. Incentivizing these activities, and embedding them in appropriate responsible institutions, will enable future refinements to risk models, and improved accuracy in model outputs. The momentum from the design and implementation of a regional risk transfer facility will further strengthen these incentives.

Notwithstanding the uncertainties and limitations detailed throughout this document, the risk modelling conducted under this TA has delivered a considerably enhanced understanding of earthquake, flood, and infectious disease risk across the region. This provides a sound basis for discussions with key stakeholders across the region, recommendations surrounding disaster risk financing, and future priorities to support increased resilience in the future.