METHODOLOGICAL NOTE:

2024 REACH - CCCM NATIONAL FLOOD HAZARD ANALYSIS OF IDP SITES IN YEMEN (MARCH 2024)

Background & Objectives

Flooding is one of the major threats to Internally Displaced Persons (IDPs) residing in IDP hosting sites across Yemen. For instance, the prevalence of flooding across Yemen is evident in the CCCM Flood Report, which listed 893 flooding events occurring between June 2021 and January 2024.[1] Meanwhile, the latest REACH-CCCM Cluster site profiling data indicated that based on perceived risk and previous occurrence, flooding was reportedly a site threat in a majority of assessed managed sites in Ansar Allah areas (AA-areas), 10% of assessed managed sites in Internationally Recognised Government (IRG) areas and just-under half of unmanaged sites in IRG-areas, indicative of heightened flood hazard exposure in AA-areas and unmanaged sites.[2] Data from managed sites in IRG-areas indicated that a considerable minority (23%, *November 2023*) lacked flood contingency planning, a figure which is likely to be higher in unmanaged sites.

This National Flood Hazard Analysis is produced by REACH, in collaboration with the Coordination and Camp Management Cluster (CCCM) Cluster and aims to triangulate the data sources on flooding in IDP sites across Yemen to classify each IDP hosting site with a 'flood hazard' score. This score will provide an indication of the extent to which flooding is an issue in each IDP hosting site across Yemen, and any patterns in variability across different geographies / site typologies. Consequently, this National Flood Hazard Analysis aims to be a supporting data source for 2024 flood contingency planning by the CCCM Cluster and OCHA. This methodological note is an updated version of the 2023 CCCM National IDP Site Flood Hazard Analysis, adjusted to reflect on the 2023 findings and lessons learned.

[1] This figure was calculated by combining flood incident reports provided to REACH by the CCCM Cluster covering reported incidents since June 2021.
[2] REACH SMT R8, SRT 2023 (North Dataset), SRT 2023 (South Dataset – Unmanaged Sites in IRG-Areas)

Scope & Data Sources

This analysis aims to provide an estimated Flood Hazard Score for all IDP sites in Yemen by triangulating all the data sources available. The data sources and their respective coverage is outlined below:

Primary data sources

- **CCCM Flood Report**: Official CCCM Flood Tracking system. REACH will consider flood data collected since the establishment of the CCCM Flood Report in 2021. CCCM Flood Report data backdated to 2021 will be considered for the purpose of this analysis due to rain pattern variations over time (at-risk sites may not flood annually.
    - 893 flooding/heavy rain events have been officially reported to the CCCM Flood Report since June 2021 (status: June 2021 – January 2024)
- **CCCM IDP Hosting Site Master List (November 2023)**: Official CCCM Master List presenting the total number of IDPs in IDP hosting sites. The latest IDP Hosting Site Master List indicated that there are 2,284 IDP hosting sites in Yemen across 22 governorates (status: December 2023)
- **CCCM Site Monitoring Tool (SMT):** Contains indicators on IDP hosting site exposure to flooding, categorized by 'very high', 'high', 'medium' or 'low' hazard exposure, in addition to an indicator on flood occurrence and perception of flooding as a site threat. This data source is based on information compiled by site managers across managed IDP sites in IRG-controlled areas. 2023 SMT data across R1-R8 covered 308 IDP hosting sites in IRG-controlled areas of Yemen.
- **CCCM Site Reporting Tool (SRT):** Contains indicators on IDP hosting site exposure to flooding, categorized by 'very high', 'high', 'medium' or 'low' hazard exposure, in addition to an indicator on flood occurrence and the perception of flooding as a site threat. The latest SRT data (April – June 2023) covered 268 managed sites in AA-controlled areas and 365 unmanaged sites in IRG-controlled areas.
- **2023 SNCC Feedback:** In 2023, Sub-National Cluster Coordinators (SNCCs) provided feedback on the 2023 Draft IDP Site Flood Hazard scores. Their feedback was based on their and partner's field knowledge, including historical flooding events. This covered 1940 IDP sites with estimated flood hazard scores. These scores are considered relevant for the 2024 analysis, but are subject to final re-confirmation from the CCCM Cluster / SNCCs.

Back-up data sources

REACH Regional Flood Hazard mapping (called HEC-RAS modelling[3]): 18 watershed basins covering 522 IDP sites across 43 districts and partially covering 7 governorates (production date: January 2022 – December 2023). **Unlike in previous years, HEC-RAS modelling will only be used in instances where IDP site information was not available in either the CCCM Flood Report, site profiling tools (SMT/SRT) and 2023 SNCC feedback.**

Table 1, outlined below, shows the coverage of IDP hosting sites across Yemen by the respective scope of each of the above data sources.

Table 1. Coverage of Flood data sources 2021-2024

Data source	Number of IDP sites reported on
2023 National IDP Site Flood Hazard Analysis	Total scores for **2301 IDP sites** Estimated flood hazard scores of SNCCs: **1940 sites**
CCCM Site Monitoring Tool (SMT)	**308 sites** covered across R1-R8 in 2023 (managed sites in IRG-areas *only)*[4]
CCCM Site Reporting Tool (SRT) 2023	**268** managed sites in AA-areas and **365** unmanaged sites in IRG-areas[5]
REACH HEC-RAS modelling (January-September 2022)	30 M: 409 sites 12 M: 42 sites 2.5 M: 74 sites **Total: 522 sites[6]**

[3] HEC-RAS stands for Hydrologic Engineering Center's - River Analysis System.
[4] There are a higher number of existing sites covered across SMT R1-R8 than the number of managed sites in IRG areas of Yemen as of the November 2023 Master List, due to the fact that SMT R1-R8 covered sites that continued to exist but were no longer covered by SMC agencies at the time of the master list publication.
[5] No data was collected for unmanaged sites in AA-areas during 2023
[6] The total excludes sites which may no longer exist and are counting under the 30M, 12M and 2.5M modelling referred to above. After filtering these sites, 522 sites were provided with a HEC-RAS score.

CCCM Flood Report (June 2021-January 2024	**892** Flood incidents since June 2021

Interpretation of Flood Hazard Scores

The final **Flood Hazard Scores** per IDP site refer primarily to the **likelihood** of a flooding event occurring in a site based on either previous flood occurrence and/or perceptions of the site's exposure to flooding. REACH's HEC-RAS Flood Hazard derived scores also refer to the *potential depth* of a flood event. In addition, SNCC derived Estimated Flood Hazard Scores might also refer to the number of people/assets affected, number of flood events happening or whether any flood prevention activities have been implemented in the site (and thereby reducing the flood hazard). The final Flood Hazard Scores could be used to **support prioritization of flood prevention activities for specific IDP sites at national level**. However, further detailed site flood hazard assessments with field partners are necessary to understand the actual extent of flood exposure based on the material conditions of the site, to in turn improve understanding of the potential impact of a flooding event and in formulating appropriate flood response plans.

NOTE: When interpreting Flood Hazard **Scores**, it is important to note **they should be considered as indicative estimates,** largely due to the limitations in the data available across the data sources. Please see **Table 2** for an overview of each data source and its limitations and **Table 3** for a clear overview of how Flood Hazard Scores were calculated to ensure the appropriate interpretation of flood hazard scores.

Analysis Framework

This analysis will differ from the 2023 Analysis Framework by favoring the determination of a Flood Hazard Score for IDPs sites based on **CCCM Flood Report** and **SMT/SRT** data, with the utilization of HEC-RAS analysis only occurring in instances where the IDP hosting site is not covered by the former data sources. Therefore, the analysis framework will be comprised of the following steps:

1. **Data Triangulation: Determination of Estimated Flood Hazard Scores for all IDP sites**: REACH will triangulate all available data sources (i.e. CCCM Flood Report, SMT, SRT, 2023 SNCC Flood estimates from National Flood Hazard Analysis) with each other to devise a 2024 Draft Flood Hazard Score per site.
2. **REACH HEC-RAS analysis (only if no data available for previous step):** Between 2021-2023, REACH mapped a total of 17 watersheds, partially covering 8 governorates. Based on this flood hazard mapping, REACH will develop flood hazard (HEC-RAS) scores for all sites that overlap with REACH's Flood Hazard Mapping.
3. **Review by CCCM National Cluster team & Sub-National Cluster Coordinators**: The CCCM Cluster and Sub-National Cluster Coordinators (SNCC) (potentially in collaboration with Area-Based Coordinators and CCCM partners) will review the Draft 2024 Flood Hazard Scores developed by REACH – and consider whether updates are required to the 2023 SNCC feedback. Following this, REACH will integrate the SNCC feedback to determine the Final Flood Hazard Scores per site.

1. Data Triangulation: Determination of Flood Hazard Scores per site

Following the Flood Hazard (HEC-RAS) modelling and scoring, REACH will determine the Draft 2024 National Flood Hazard Scores for all IDP sites based on the triangulation of REACH-CCCM SMT/SRT tools, the CCCM Flood Report and 2023 SNCC Flood Estimates. The HEC-RAS Flood Hazard Severity Scores will only be utilized to estimate the flood hazard of IDP hosting sites in instances where information is unavailable in either the CCCM Flood Report or the REACH-CCCM site profiling (SMT/SRT) tools.

2. Determination of HEC-RAS Flood Hazard scores for sites not covered by data triangulation

For sites that are not covered by the above data sources, REACH will determine specific Flood Hazard (HEC-RAS) scores per IDP site, which are based on flood hazard models developed by REACH and UNOSAT between 2021-2023. Based on this analysis, flood hazard and depth models should be available for a total of 525 sites. In general, REACH will overlay the exact location of IDP sites with the available flood hazard models and based on an estimated buffer zone around the IDP site determine the HEC-RAS Flood Hazard scores.[7] See **Annex 1** for detailed technical methodology on REACH's flood hazard mapping.

To properly interpret the Flood Hazard Scores, it is important to note that each data source has a different methodology and limitations, as described in **Table 2**.

Table 2. Overview of methods and limitations of data sources

Data source	Method	Limitations
REACH-CCCM Site Monitoring Tool (SMT) R1-R8 (January-November 2023)	**CCCM Site Monitoring Tool (SMT) contains** an indicator on flood exposure of an IDP site, categorized across 'very high'. 'high', 'medium' and/or 'low' thresholds, with a relevancy constraint based on reporting 'flooding' and/or 'heavy rain' as a site threat	This data source reflects the perceptions of the site manager based on the data sources available to them. This is not based on dedicated spatial flood hazard criteria and analysis.
REACH-CCCM Site Reporting Tool (SRT) (April - June 2023)	**CCCM Site Reporting Tool (SRT) contains** an indicator on flood exposure of an IDP site, categorized across 'very high'. 'high', 'medium' and/or 'low' thresholds, with a relevancy constraint based on reporting 'flooding' and/or 'heavy rain' as a site threat	This data relies on reports from Key Informants on whether flooding may be a threat to the site. No official flood hazard assessment. Data coverage is less comprehensive than the SMT data, covering April-June 2023. Therefore, some flood incidents that occurred during the remainder of the 2023 flooding season may have gone unreported.
CCCM Flood Report (June 2021-January 2024)	**CCCM Flood Report** highlights sites where flooding occurred since 2021. **Overall, the CCCM Flood Report is considered the most authoritative dataset in this analysis, since it reports actual events.** The CCCM Flood Report also allows REACH to validate its HEC-RAS model findings over time.	This dataset includes sites where reports of flooding have occurred since June 2021. It is also possible that not all flood events were flagged and/or accurately reported (particularly for unmanaged sites), so total number of flooding might be higher.
REACH HEC-RAS modelling **Only utilized in cases where no information is available from the above sources.** (January 2022 – December 2023)	**Flood Hazard (HEC-RAS) models** can provide flood hazard and flood depth products based on a *designed* storm. These products are overlaid with IDP site location and an estimated buffer / boundary to derive estimated flood hazard scores.	Based on modelling a designed storm and not an actual flooding event. Also, exact IDP site extents/boundaries are not available, and coordinates might be inaccurate. Thus, the models may overestimate or underestimate Flood Hazard.

[7] A total of 4 IDP sites were excluded that bordered REACH's Flood hazard models and no clear flood score could be determined.

SNCC Flood Estimates (February 2023)	In 2023, SNCCs provided feedback on the 2023 Draft IDP Site Flood Hazard scores. Their feedback was based on their field knowledge, including of historical flooding events. 2023 scores will be considered in 2024 – with SNCCs responsible for updating scores in cases of changes in site severity.	SNCC feedback is not available for all sites and subjective interpretation of flood hazard. This is not a substitute for an official, dedicated flood hazard assessment.

Overall, the 2024 Flood Hazard Scores for each IDP site will be developed based on the triggering of scenario that indicates high/medium/low hazard respectively (see Table 3). For 2024, a new 'critical hazard' category was added in order to differentiate the sites with the highest severity in light of the large proportion of sites that fell under the high hazard category in the preliminary scoring.

<u>The highest triggered score will be used for all sites</u>, meaning that if a site is determined to fulfill **one** of the scenarios indicative of a 'high' flood hazard score – the hazard score will be classified as 'high' irrespective of whether that site also fulfills scenarios that would classify it in the 'medium' or 'low' categories. Likewise, if a site fulfills the criteria for one 'medium' and one 'low' scenario, that site will be classified with a 'medium' site flood hazard score. The 'unknown' hazard score will be attributed to sites that lack the information available in order to make an indicative decision regarding their most-fitting flood hazard score category. In cases of **contradictory information** where the CCCM Flood Incident Report confirms a flood occurance in a site but the REACH SMT/SRT does not or vice versa, the data source that indicated a flood hazard will always be prioritised.

Table 3. Determination of Flood Hazard of IDP sites

Scenarios	IF	ESTIMATED Flood Hazard of IDP site				
		Critical hazard	High hazard	Medium hazard	No/low hazard	Unknown
Scenario A	If SMT/SRT data indicated 'very high' exposure to flooding and at-least one flood occurred according to REACH-CCCM SMT/SRT or CCCM Flood Incident Report	x				
Scenario B	At-least 5 flooding incidents recorded in the CCCM Flood Incident Report since June 2021	x				
Scenario C	Flood occurrence reported for IDP site in CCCM Flood Report **and/or** REACH-CCCM SMT/SRT[1]		x			
Scenario D	2023 REACH-CCCM SMT or SRT data indicates that the exposure to flooding is *'very high'* or *'high'*		x			
Scenario E	2023 REACH-CCCM SMT or SRT data indicates that the exposure to flooding is *'medium'*			x		
Scenario F	2023 REACH-CCCM SMT or SRT data indicates that flooding and/or heavy rain was not reported as a site threat and no flooding occurred				x	
Scenario G	In the 2023 National Flood Hazard Analysis feedback, the Sub-National Cluster Coordinator (SNCC) considered the site to fitting of the 'high' hazard clarification		x			
Scenario H	In the 2023 National Flood Hazard Analysis feedback, the Sub-National Cluster Coordinator (SNCC) considered the site to fitting of the 'medium' hazard clarification			x		
Scenario I	In the 2023 National Flood Hazard Analysis feedback, the Sub-National Cluster Coordinator (SNCC) considered the site to fitting of the 'low' hazard clarification				x	
Scenario J	Site has no available data in CCCM Flood Incident Report, REACH-CCCM Site Profiling Tools (SMT & SRT) or 2023 SNCC feedback and a HEC-RAS severity score of 3 or higher.	-	x			
Scenario K	Site has no available data in CCCM Flood Incident Report, REACH-CCCM Site Profiling Tools (SMT & SRT) or 2023 SNCC feedback and a HEC-RAS severity score of 2.	-		x		

Scenario L	Site has no available data in CCCM Flood Incident Report, REACH-CCCM Site Profiling Tools (SMT & SRT) or 2023 SNCC feedback and a HEC-RAS severity score of 1.	-			x	
Scenario M	Site has no available data in CCCM Flood Incident Report, REACH-CCCM Site Profiling Tools (SMT & SRT) or 2023 SNCC feedback and no HEC-RAS data available.					x

3. Field validation by CCCM Cluster & Sub-National Cluster Coordinators

Following the Data Triangulation and REACH's estimate on which IDP sites are at hazard of flooding, the CCCM National Cluster Coordination team, with support from the CCCM Sub-National Cluster Coordinators (SNCCs), will review the findings. This is especially relevant for sites for which REACH's provisional site score was 'unknown', cases where only HEC-RAS data was available and instances where SNCC feedback is likely to differ from 2023 estimates. SNCCs may reach out to CCCM Area Coordinators or CCCM partners in the field to verify the results.

Ultimately, REACH will integrate any further feedback from the CCCM Cluster and SNCCs to derive the Final 2024 Flood Hazard Scores per IDP site. In addition, with support from CCCM partners the CCCM Cluster will identify the total number of IDP sites that will be targeted for flood prevention and response activities based on REACH's analysis.

Outputs

Outputs can be found via **IMPACT's Resource Centre** on the dedicated REACH Yemen page:

- **Dataset** containing suggested 2024 Flood Hazard Scores per IDP site
- **National map** highlighting Final 2024 Flood Hazard Scores per IDP site
- **National map** illustrating historical flood events in affected IDP sites from June 2021- January 2024
- **Regional HEC-RAS maps** highlighting total 2024 Flood Hazard Scores per IDP sites (Hajjah, Al Hodeidah, Taiz, Marib)
- **Presentation** of findings to CCCM Cluster, Inter-Cluster Coordination Group (ICCG) and partners

Limitations

As part of this analysis, REACH identified the below limitations for the analysis in general, and the HEC-RAS flood hazard modelling limitations specifically can be found in **Annex 1**:

Information gaps: All data sources from which flood hazard scores can be derived have coverage data gaps (i.e., HEC-RAS modelling, CCCM SRT/SMT and Flood Report). Based on these datasets, Estimated Flood Hazard Scores could be derived for 1840 sites out of 2285 sites (81%).

Contradictory information: All data sources have contradictory information over time and between each other. This could stem from reporting errors, actual changes over time or the fact that each dataset has a different methodology and limitations. In cases where primary data sources provided contradixtory data on site flood hazards, the data which suggested the highest flood risk is always used.

Different methodologies of data sources: All data sources have a very different methodologies and results in terms of deriving the potential flood hazard per site, with each of them having unique limitations.

Interpretation of results: Based on the limitations noted in this section, Flood Hazard scores should be considered as indicative estimates.

Not a comprehensive flood risk analysis: This analysis focuses on the site exposure to flooding and presence of recent flooding events (since June 2021) – while this encompasses flood exposure at site-level, it does not analyse the flood exposure or vulnerability of specific households or shelters based on their typology.

Site boundaries: The exact site extent/site boundaries of IDP sites in Yemen are not available to this date. Thus, REACH/CCCM had to develop an estimate of the extent of each IDP site used in the HEC-RAS analysis. This was done by establishing estimated buffer radiuses based on population size, which may not be accurate since IDP population density per site is also not known (see section: Phase 1. HEC-RAS Analysis – Buffer radius).

Site location: Exact site location is only available for 1,153 out of 2,302 sites. In addition, not all site locations have been verified, so they might contain errors.

HEC-RAS score: HEC-RAS modelling might slightly overestimate or underestimate flood hazard, due to above mentioned limitations and the fact that REACH's developed scoring system is not an officially hydrologically tested system.

Annex 1. REACH Flood Hazard Mapping (HEC-RAS modelling) – Technical Note

Background & Objectives

To map flood hazard across Yemen, REACH modelled 18 separate watershed basins through Hydrologic Engineering Center's - River Analysis System (HEC-RAS). A two-dimensional (2D) unsteady flow hydraulic model was built using HEC-RAS to enable deriving flood hazard and depth products. The results from these types of modeling outputs can provide a high-level understanding of flood hazards on a catchment-wide scale and help to identify flood susceptible areas, especially areas at hazard of flash flooding.[8] These products can be overlaid with any relevant spatial information (i.e., IDP sites, cities, critical infrastructure) to understand its hazard of flooding. For this specific national assessment, REACH overlaid the HEC-RAS models with IDP site locations to derive estimated flood hazard scores.

Coverage

REACH conducted flood hazard (HEC-RAS) models covering **18 basins** that partially cover 7 governorates, as described below.

- **30 M DEM:** 13 basins across Hajjah. Al-Hodeidah, Taiz, Al-Mahwait, Raymah, Dhamar, Sana'a and Marib

- 12 M DEM: 1 basin in Abs, Hajjah

- 2.5 M DEM: 4 basins in Abs, Hajjah & Marib

Catchment areas with a higher overall number of IDP population and IDP population density were prioritized for this initial phase of the exercise, which included primarily basins from the Hajjah, Al Hodeida, Taiz and Marib governorates. HEC-RAS 2D modeling is a detailed and computationally intensive process, which makes it a time-consuming task. Due to its complexity and the amount of data required, it is not possible to model the country at once. It would require a lot of computational power, a vast amount of data and a lot of time. Instead, the model is typically applied to specific areas and watersheds.

[8] Flooding that begins within 6 hours, and often within 3 hours, of the heavy rainfall (or other cause). Flash Floods can be caused by a number of things, but is most often due to extremely heavy rainfall from thunderstorms. Flash Floods can occur due to Dam or Levee Breaks, and/or Mudslides.

Input data

Precipitation data

The model output products inform the extent, depth, and hazard of areas where flooding may occur based on extreme precipitation events of a 25-year return period. The model applies a direct precipitation method, where precipitation is applied to all cells generated in a computational mesh. Due to the lack of trustworthy precipitation data from weather stations in Yemen, data from Saudi Arabia was used in basins in the Hajjah governorate, near the northern border, and for all other basins modeled in this analysis, precipitation data derived from satellite imagery was used.

The precipitation data from Saudi Arabia is from the Abha city region, which is a coastal wadi area with similar characteristics as the catchments found in Hajjah, located approximately 100 km north of the Yemeni border in Hajjah.[9] The Saudi precipitation time-series data was processed into intensity-duration-frequencies (IDF) to generate rainfall temporal distribution data. The IDF data was then used to design a 12-hour storm with 10-minute steps for the 25-year return period using the alternating block method.[10] The storm data was validated using Google Earth Engine (GEE) by analyzing a real-life storm that occurred in the same coastal region in Yemen, bordering Saudi Arabia. The storms displayed very similar durations and total depths as the designed storm.

The satellite precipitation data used in the modeling is from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) dataset, which builds on 'smart' interpolation techniques and record precipitation estimates based on infrared Cold Cloud Duration (CCD) observations, incorporating station data and the spatial correlation structure of CCD-estimates in interpolation methods.[11] The CHIRPS data is processed through a series of statistical techniques to estimate extreme precipitation based on the Generalized Extreme Value (GEV), a commonly used statistical distribution in precipitation analysis. The 25-year return precipitation values were disaggregated into 12-hour storm events based on sub-daily ratios of precipitation, also taken from the Jeddah region in Saudi Arabia, with similar characteristics as the Yemeni Coastal area.[12] Ratios for mountainous and transitional areas will also be used in future modelling analyses.

[9] Al-anazi, K. and El-sebaie, I. (2013) Development of Intensity-Duration-Frequency Relationships for Abha City in Saudi Arabia. International Journal of Computational Engineering Research, 3, 58-65.
[10] A designed storm is an artificial hyetograph that takes the precipitation depths for time intervals over a specified total storm duration.
[11] Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, James Rowland, Laura Harrison, Andrew Hoell & Joel Michaelsen. "The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes". Scientific Data 2, 150066. doi:10.1038/sdata.2015.66 2015.
[12] Awadallah, A. (2015), Regional I-D-F for Jeddah, Saudi Arabia. J. Flood Hazard Manage, 8: 195-207. https://doi.org/10.1111/jfr3.12085

Terrain and infiltration data

The terrain data used to generate the computational mesh of the 2D flow areas for this analysis was as following:

1. **High-resolution 2.5 and 12** meters digital elevation models (DEMs) were used in 5 basins in Abs, Hajjah and Marib.
2. **A 30-meter** resolution digital elevation model (DEM) product called AW3D30.[13] This product is available freely online thanks to an open license conceded by the Japan Aerospace Exploration Agency (JAXA), and it was used to model the other basins across Hajjah, Al Hodeida, Taiz, Al Mahwit, Raymah, Dhamar, Sana'a and Marib.

The 2D flow area extents were defined for each basin in the HEC-RAS processing environment, and computation points were generated. The datasets used to complement the analysis include landcover data[14] from the European Space Agency (ESA) and hydrological soil group (HSG) data.[15] An infiltration layer is created based on these two products, to account for the precipitation losses by infiltration and help determine the excess or net rainfall. For this purpose, the Soil Conservation Service (SCS) Curve Number (CN) approach is adopted, where estimates of the CN are a function of the different combinations of hydrologic soil groups and land cover types present in the terrain.[16] The CNs are derived from reference tables suggested in the HEC-RAS documentation. The landcover product has 11 different categories which are all assigned a roughness coefficient (Manning's value), also based on HEC-RAS documentation.

Development of Estimated HEC-RAS Flood Hazard Scores

To develop HEC-RAs flood hazard scores per IDP site, REACH had to follow the below steps:

1) Design buffer radius per IDP site, since exact IDP site extents are not available

2) Overlay IDP site location with REACH 2.5M, 12M and 30M Flood Hazard & Depth products

3) Develop estimated HEC-RAS Flood Coverage Severity Scoring

[13] Japan Aerospace Exploration Agency (2021). ALOS World 3D 30 meter DEM. V3.2, Jan 2021. Distributed by OpenTopography. https://doi.org/10.5069/G94M92HB. Accessed: 2022-03-03.
[14] Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, Linlin, Tsendbazar, N.E., Ramoino, F., Arino, O., 2021. ESA WorldCover 10 m 2020 v100. https://doi.org/10.5281/zenodo.5571936.
[15] Ross, C.W., L. Prihodko, J.Y. Anchang, S.S. Kumar, W. Ji, and N.P. Hanan. 2018. Global Hydrologic Soil Groups (HYSOGs250m) for Curve Number-Based Runoff Modeling. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1566.
[16] SCS. 1956. Hydrology. National Engineering Handbook, Supplement A, Section 4. Soil Conservation Service, US Department of Agriculture: Washington, DC; Chapter 10.

Buffer radius

Information regarding the extent or boundaries of IDP hosting sites in Yemen is currently unavailable. Accordingly, REACH had to establish an estimated buffer radius for each IDP hosting site. By utilizing boundary maps of 32 IDP sites, which were further divided into multiple segments for each site, we developed a machine learning algorithm using the Orange Data Mining tool to estimate the buffer radius based on population density. The analysis revealed that the population range and buffer radius trends did not follow a distinct pattern; instead, exhibited a higher level of precision in determining buffer sizes. In our previous methodology, buffer sizes were determined based on a range set by REACH Yemen. The following outlines the methodology used in our previous approach.

Information on the extent or boundaries of IDP hosting sites is not available in Yemen. Accordingly, REACH had to establish estimated buffer radiuses for each IDP hosting site. Based on the below two example CCCM Site Plans, REACH created **different estimated buffer radiuses based on IDP population size**. See Table 1 below for more details.

Table 1. Buffer radius of IDP hosting sites

	Very small sites	Small sites	Medium sites	Large sites	Very large sites
Population size	1-200	201-500	501-1,000	1,001-5,000	>5,000
Buffer radius	100m	200 m	300 m	500m	1km

The below two example CCCM Site Plans were used to estimate the site boundaries in Table 1.

Site A: 80 HHs (around 480 people): 47 M radius

Site B: 512 HHs (around 3072 people): 303 M radius

Step 1: Classify the flood hazard percentage coverage (%) and flood depth coverage (%) per **severity class** (0-5, see Table 2) through ArcGIS.

Step 2: **Aggregate** the flood hazard coverage (%) and flood depth (%) per buffered IDP site to derive the **severity class** per indicator (see Table 3).

Step 3: Determine the **final severity scores per IDP** sites based on the **highest value** of the two indicators.

HEC-RAS Flood Coverage Severity Scoring

In the absence of IDP site data from the CCCM Flood Report and/or SMT/SRT data, IDP site hazard scores will be calculated using HEC-REAS Flood Coverage (where available). For the **HEC-RAS model Flood Hazard Severity Scoring**, REACH will use two main indicators, which are based on **flood hazard and flood depth** (see Table 2 below). Based on these two indicators, REACH will implement the below **steps** to draw up a HEC-RAS Flood Hazard Severity Score per IDP site:

Step 1: Classify the flood hazard percentage coverage (%) and flood depth coverage (%) per **severity class** (0-5, see Table 2) through ArcGIS.

Step 2: **Aggregate** the flood hazard coverage (%) and flood depth (%) per buffered IDP site to derive the **severity class** per indicator (see Table 3).

Step 3: Determine the **final severity scores per IDP** sites based on the **highest value** of the two indicators.

Table 2. HEC-RAS Flood Hazard Severity Classes per Indicator

The below table shows the severity classes for flood hazard and flood depth, which are used in cases where the site is lacking information from 2023 SNCC feedback, SMT/SRT 2023 data and incidents listed in the 2023 CCCM Cluster Flood Incident Report. It was decided that **site severity scores that are based on HEC-RAS data cannot be categorized as 'critical hazard'** due to the limitations of this data source. Hence, scores are capped at 'high hazard' and sites with scores between 3-5 will be classified accordingly.

Nr	Indicator	0 – No value (no hazard)	1 – Low hazard	2 – Medium hazard	3 – High hazard	4 – Very High hazard	5 – Extreme hazard
1	flood hazard[17] (m^2/s)	No value	≤ 0.2	>0.2 – 0.5	>0.5 – 1.5	>1.5 – 2.5	>2.5
2	flood depth (m)	No value	≤ 0.5	>0.5 – 1	>1.0 – 2.0	>2.0 – 5.0	>5.0

[17] Flood Hazard is defined as = flood velocity x flood depth

Table 3. Determination of HEC-RAS Flood Hazard Score per IDP site

The below table will be applied to both flood hazard and flood depth indicators.

No hazard	Low hazard	Medium Hazard	High Hazard		
0 – No value	**1 – Low**	**2 – Medium**	**3 – High**	**4 – Very High**	**5 – Extreme**
No hazard/no value (for 2.5 and 12 M resolution)	All remaining categories **OR** No hazard/no value (for 30 M resolution)[18]	If Severity classes 2, 3, 4 and 5 are >=10% of total buffer site area. **OR** If Severity classes 2, 3, 4 are >=20% of total buffer site area. **OR** If Severity classes 2 and 3 are >=30% of total buffer site area. **OR** If Severity class 2 is >=40% of total buffer site area.	If Severity classes 3, 4 and 5 are >=15% of total buffered site area. **OR** If Severity classes 3 and 4 are >=20% of total buffered site area. **OR** If Severity class 3 is >=30% of total buffered site area.	If Severity classes 4 and 5 are >=15% of total buffered site area. **OR** If Severity class 4 is >=20% of total buffered site area.	If Severity class 5 is >=20% of total buffered site area.

[18] Since the HEC-RAS model with 30M resolution is less reliable, IDP sites with no hazard scores are still placed under severity class "Low Hazard". Contrary, for IDP sites with 2.5 and 12M resolution a "No Hazard" score is assigned, since results are considered more reliable.

Example: HEC-RAS Flood Severity Score Calculation per IDP site

Step 1: Classification of flood hazard (%) and flood depth (%) indicator per severity class

REACH will classify the flood hazard percentage coverage (%) and flood depth coverage (%) per **severity class** (No hazard – Extreme hazard, see Table 2) through ArcGIS. *This classification is done automatically in ArcGIS.*

Nr	Indicator	0 – No value (no hazard)	1 – Low	2 – Medium	3 – High	4 – Very High	5 – Extreme
1	% of covered IDP site area by flood hazard	35%	0%	0%	10%	50%	5%
2	% of covered IDP site area by flood depth	35%	45%	15%	5%	0%	0%

Step 2: Aggregation of flood hazard (%) and flood depth (%) per severity class

REACH will aggregate the flood hazard coverage (%) and flood depth (%) per buffered IDP site to derive the **severity class** per indicator. *This aggregation is calculated in Excel, based on the ArcGIS values.*

Indicator	0 – No value (no hazard)	1 – Low	2 – Medium	3 – High	4 – Very High	5 – Extreme	Aggregation
% of covered IDP site area by flood hazard	35%	0%	0%	10%	50%	5%	55%

Step 3. Calculate Total Severity Score per IDP site

REACH will determine the **final severity scores per IDP** sites based on the **highest value** of the two indicators. *This is calculated in Excel.*

Example

Indicator 1: Severity Classes 4 and 5 = 55% = **Severity Score 4**
Indicator 2: Severity Classes 2, 3, 4 and 5 = 20% = **Severity Score 2**

Total severity Score = MAX of Indicator 1 and Indicator 2 = 4

Technical Limitations

In addition to the analytical limitations established, there are also the following technical limitations with the HEC-RAS analysis:

- **Technical limitations**

    - **30 M terrain data:** A 30-meter DEM resolution was used for 13 basins out of 17 basins in the modelling, which likely underestimates the depth of smaller water flows and overestimates the extent.
    - **Hydraulic structures**: Bridges and culverts, piped drainage networks, irrigation canals and open channels have not been included in the hydraulic model.
    - **Precipitation**: The precipitation data used in different governorates are from different sources, and although they have been processed using scientifically established methods, they may differ significantly and there is currently no way to validate the data.
    - **Storm-events**: The storm events designed to be used as inputs for the models varied in duration (and intensity) due to the different methods used to disaggregate the data. Storms of 6 to 12-hour durations were used in our analysis, however flooding events can sometimes be caused by several consecutive storm events that can last for several days. This phenomenon was not modelled in this analysis.