

# UKRAINE: SEVERITY MODEL METHODOLOGY NOTE

March 2024

## TABLE OF CONTENTS

---

Introduction.....	3
About the ACAPS Ukraine Severity Model .....	3
Rationale .....	3
Aim .....	3
Concept of severity in the ACAPS Ukraine Severity Model .....	3
Development process.....	3
Scope and scale of the model .....	4
Model framework .....	4
Variable selection.....	4
Normalisation.....	5
Aggregation .....	9
Weighting.....	10
Final score structure.....	10
Statistical analysis .....	11
Correlation matrix of dimensions.....	11
Correlation matrix of indicators.....	11
Robustness and sensitivity analysis .....	13
Limitations .....	16
Repeated indicators.....	16
Data limitations: frequency of update.....	16
Other data limitations .....	16
Annex I. How to use the index: a hands-on guide.....	17
Getting the data.....	17
Annex II. Transformations .....	21
Sources .....	21

## INTRODUCTION

---

The full-scale escalation of the conflict between Russia and Ukraine has passed the two-year mark. The overall situation, movement of people, number of conflict-related events, and needs keep changing per context.

Assessing the crisis severity throughout time in each geographical location is crucial for the analysis of the humanitarian situation.

The ACAPS Ukraine Severity Model is a publicly available, reproducible model for assessing the crisis severity in each Ukrainian oblast per month. Input data in the model comes from numerous publicly available sources, which differ in terms of time and space granularity, scale, representativeness, and update frequency. The data is available in the [ACAPS Ukraine Master Dataset](#), itself a consolidated collection of publicly available sources, and the [ACAPS Ukraine Subnational Access Score](#).

This methodological note explains the process behind the design, structure, and limitations of the ACAPS Ukraine Severity Model, as well as the statistical analysis performed.

## ABOUT THE ACAPS UKRAINE SEVERITY MODEL

---

### Rationale

Several challenges persist in the data analysis landscape related to the Ukraine crisis. A large amount of data goes unanalysed, representing a missed opportunity for valuable insights. Without a relevant tool to consolidate and contextualise trend analysis, patterns can go unnoticed, limiting the ability to make informed decisions. The sector will also strongly benefit from a tool capable of regularly tracking the overall severity of the humanitarian situation, effectively consolidating information from different sources.

### Aim

The primary goal of this work is to develop a severity model at the oblast level (admin1), focusing on tracking key contextual, humanitarian, and conflict-related indicators. This model will enhance the understanding of the severity of the humanitarian situation in Ukraine.

### Concept of severity in the ACAPS Ukraine Severity Model

The ACAPS Ukraine Severity Model aligns its definition of severity with the [INFORM Severity Index concept note](#). While the two methodologies differ in context and data availability, the overall definition of severity remains:

“Severity of humanitarian crisis is an extent of outcomes, in terms of distribution of people affected by the severity of conditions they are confronted within the crisis, generated by the impact of a crisis and worsen by how complex is to deliver humanitarian response in the operational environment [sic].

Measure of crisis severity summarises a wide range of quantitative information about severity of humanitarian crisis and presents it in numerical and/or categorical scale to be easily used in decision making.” (EC 2020)

### Development process

The development of the ACAPS Ukraine Severity Model began during a three-day workshop organised by ACAPS, from 24–26 May 2023. Several ACAPS analysts and external stakeholders were invited to discuss, share thoughts, engage in hands-on activities, and conduct preliminary data analysis. The workshop focused on the different phases of the construction of the index: the theoretical framework, input data, and aggregation techniques.

The discussion of the theoretical framework allowed for the clarification of the concept of severity and the identification of a nested structure for its various subgroups.

In parallel, a data review was conducted to identify which indicators were available in existing datasets and provide the granularity in time and space needed for this index. ACAPS is already releasing the ACAPS Ukraine Master Dataset monthly, which consolidates key indicators of the Ukraine crisis to compare and analyse available data and track changes in the humanitarian situation, sourcing data from publicly available information.

The findings from the workshop were used as a starting point to further improve the model. After numerous discussions on dimensions, new data inclusion, aggregation, and weighting, and intermediate results, the index was finalised into its current structure.

## Scope and scale of the model

The spatial scope of the Ukraine Severity Model is subnational, with a score provided for each oblast (admin1 level).

The score is updated monthly to reflect changes in the input data. Not all the input data changes every month. When there are no updates, the most recently available data is used. The regular releases reflect the changes in the regions of Ukraine presented in the updates in the ACAPS Ukraine Master Dataset, the ACLED dataset, the ACAPS Damages dataset, and the ACAPS Ukraine Subnational Access Score.

## MODEL FRAMEWORK

### Variable selection

The ACAPS Ukraine Severity Model is based on data from multiple sources: the ACAPS Ukraine Master Dataset, the ACAPS Ukraine Subnational Access Score, the ACAPS Ukraine Damages dataset, and the ACLED dataset. When selecting the input data, we used the following criteria:

publicly available, reliable source

at least admin1 spatial granularity

updated at least four times per year.

Exceptions to the third point were made for sources that did not have updates at least four times per year but were considered essential to understanding severity, notably the number of people in need from the Humanitarian Needs Overview.

Table 1. Full list of the input variables

VARIABLE NAME	SOURCES	USED IN ACAPS DATASET	NOTES
# fatalities	ACLED	Yes (ACAPS Master Dataset)	
# battles	ACLED	Yes (ACAPS Master Dataset)	
# explosions/remote violence	ACLED	Yes (ACAPS Master Dataset)	
# km <sup>2</sup> controlled by Russian forces	ISWN; Live UA Map; ISW	Yes (ACAPS Master Dataset)	
# km <sup>2</sup> landmass affected	OCHA	Yes (ACAPS Master Dataset)	
# km <sup>2</sup> unconfirmed control	ISWN; Live UA Map; ISW	Yes (ACAPS Master Dataset)	
# people in need	OCHA	Yes (ACAPS Master Dataset)	
# people in (1–5) humanitarian condition	OCHA	Yes (ACAPS Master Dataset)	
# total infrastructure damage	ACAPS Damages dataset	Yes (ACAPS Master Dataset, ACAPS Ukraine Subnational Access score)	Used partially in Subindicators 7.2 and 9.2 of the ACAPS Ukraine Subnational Access Score
# registered IDPs	IOM	Yes (ACAPS Master Dataset)	
# total population	SSS; UNFPA; UNHCR; IOM	Yes (ACAPS Master Dataset)	
# people affected	OCHA	Yes (ACAPS Master Dataset)	
ACCESS	ACAPS Access Events dataset, ACLED, ACAPS Damages dataset	-	ACAPS Ukraine Subnational Access Score

## Normalisation

As the indicators were on different scales, a normalisation procedure was applied that, taking into account outliers and the skewness<sup>1</sup> of the data, brought all of the indicators to the same scale.

For each indicator, the distribution of the values was studied in detail, considering the data available for all of the oblasts and several months of data collection. In particular, the 0.05 and 0.95 quantiles were identified, as well as potential outliers, and the shape of the distribution was visualised to check for skewness.

The parameters for normalisation were derived from the data collected from March 2022 to November 2023 for all oblasts. While future data could slightly shift the quantiles of the distribution, these parameters were fixed to ensure the stability of the score.

For the indicators that were positively skewed, a log transformation was applied, which emphasised relative changes or differences in the order of the magnitude of values greater than the changes in absolute values (so, for instance, this transformation will result in giving more importance to a change from 500 to 1,000 than from 10,000 to 10,500).

Hence, when a logarithmic transformation was applied, the minimum for the clipping had to be larger than zero; so, for the cases where there was a lot of zeros/empty rows, the result was 1. More on this can be found in Annex 1.

To minimise the impact of outliers and to level out very high scores that did not change the situation humanitarian-wise, the indicators were 'clipped'. This means that, in the normalisation process, 0.05 and 0.95 quantiles were usually chosen as the clip minimum and maximum, respectively. Values in the dataset above or below this range were set to this clip maximum and minimum, respectively.

In most cases, the 0.05 quantile was zero. The maximum was set starting indicatively from the 0.95 quantile but considering the context of the variable, its reliability, and the impacts of this choice on the final score. More details are provided in the Case-Specific Distributions section below.

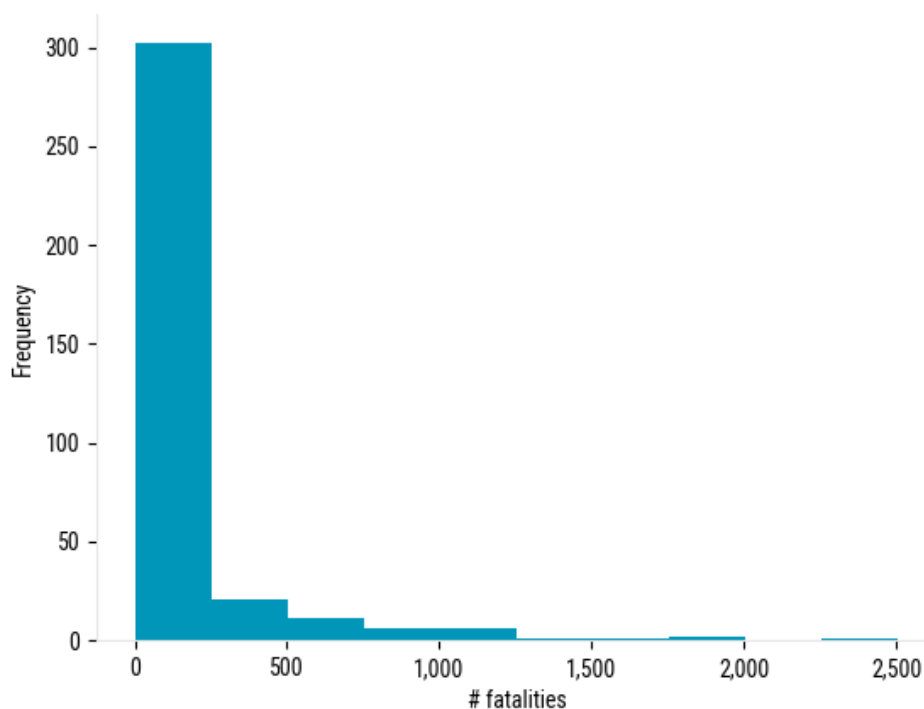
After the selection of the minimum and maximum parameters and, where needed, the application of the log transformation, the indicators were translated into a range of 0–5 using min-max normalisation. The formula for obtaining the final score of the variable is:

$$5 \cdot \frac{\text{max-value} - \text{value}}{\text{max} - \text{min}}$$

### Example of a normalisation process

An example of the standard normalisation process is presented here for the 'Number of fatalities' indicator, whose distribution is shown in Figure 1.

Figure 1. 'Number of fatalities' distribution

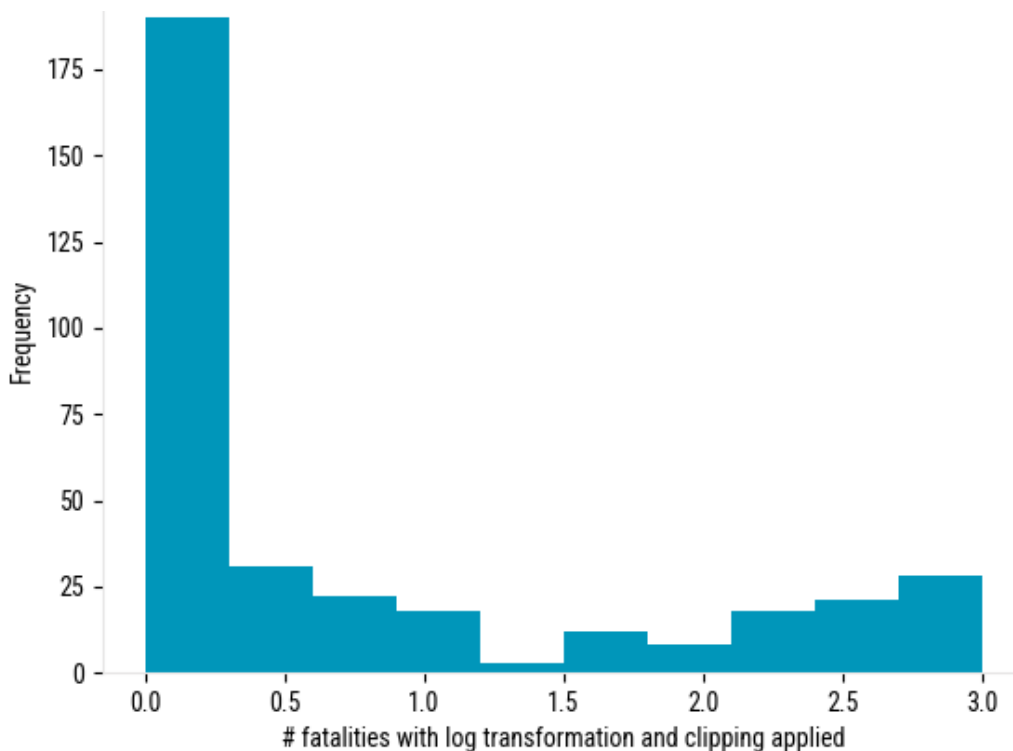


<sup>1</sup> Skewness: a measure of asymmetry within a distribution.

As seen in the distribution, the majority of data points were skewed towards lower values, so a log transformation was applied. The 0.95 and 0.05 quantiles were 726 and 0, respectively, and they were rounded (and applied to the variable) to 1,000, 0.

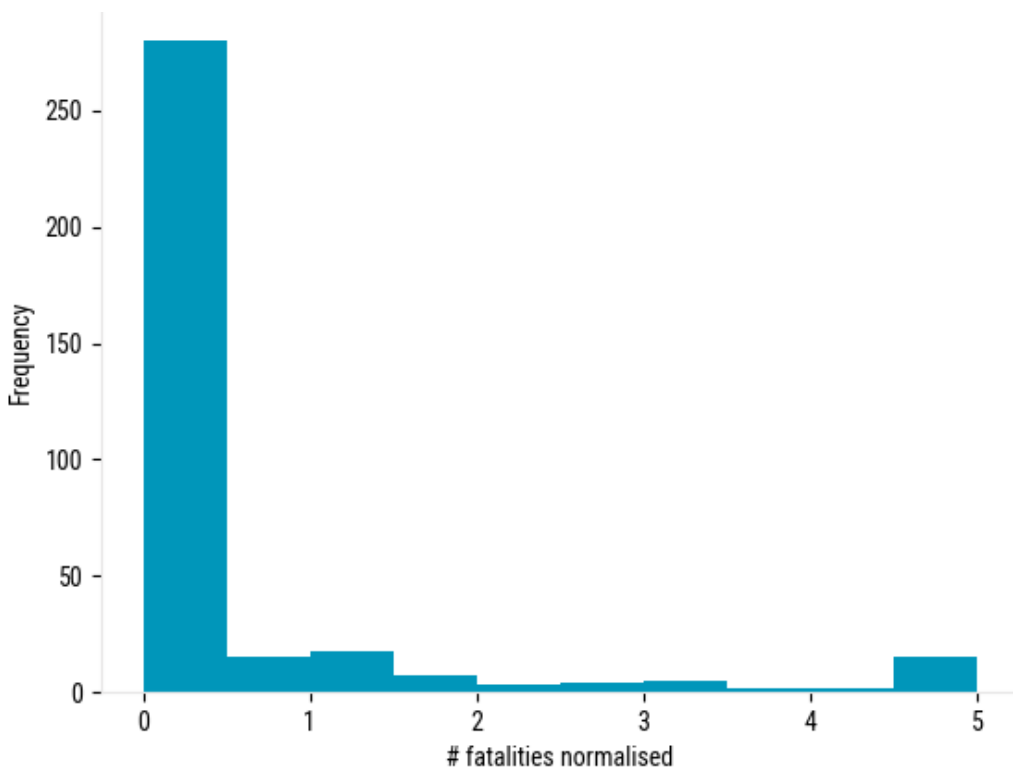
The distribution after the log transformation is shown in Figure 2.

Figure 2. The distribution of the 'Number of fatalities' indicator after clipping and log transformation



Finally, the values were put on a 0–5 standard scale using the min-max normalisation, giving the distribution shown in Figure 3.

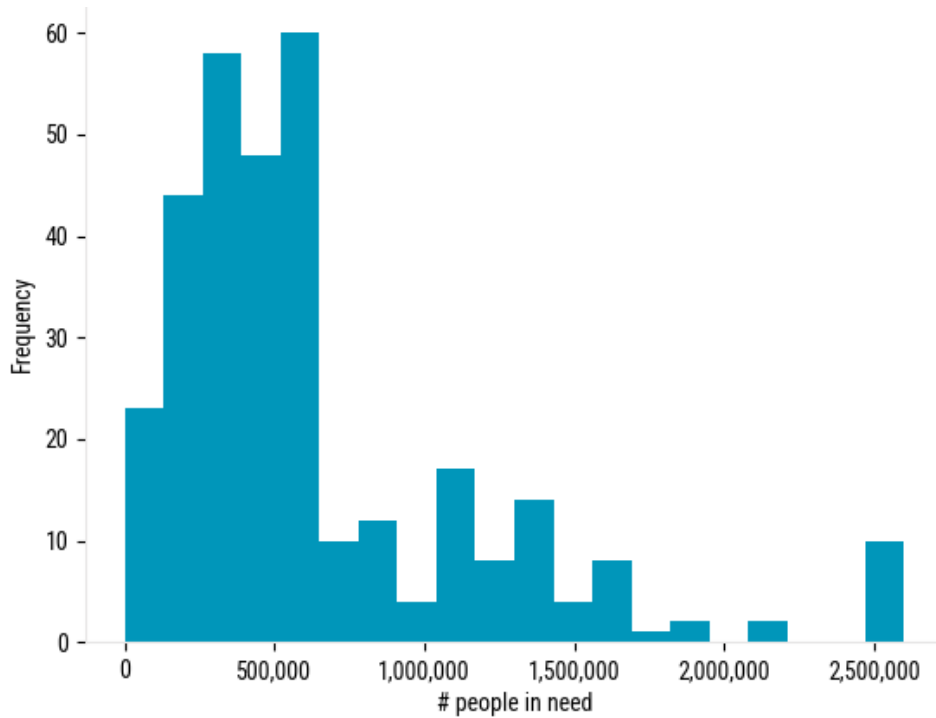
Figure 3. 'Number of fatalities' distribution after normalisation



## Case-specific distributions

The default approach to using the 0.05 and 0.95 quantiles as minimum and maximum values in the normalisation procedure was modified in cases where the higher values were not considered outliers and/or the data was not fully available and presented several empty records (zeroes). In these cases, the distribution was studied, and different approaches were tested to examine the impact on the result.

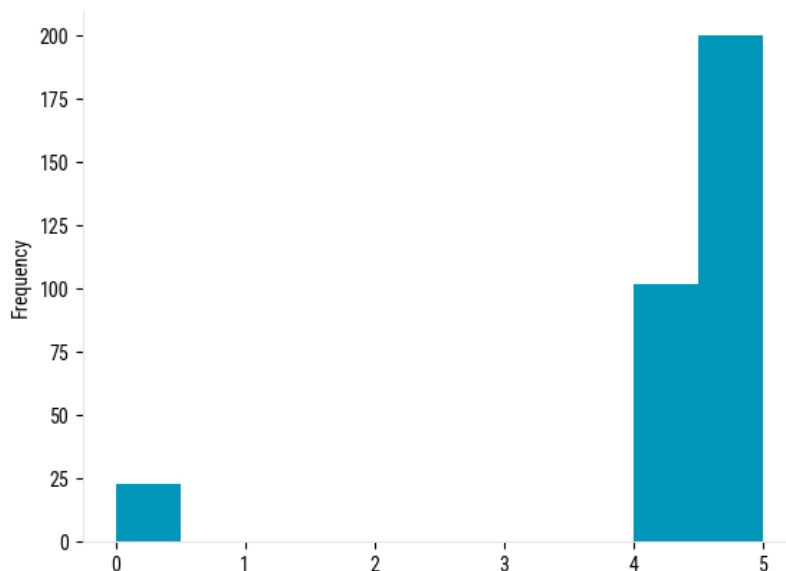
Figure 4. '# people in need' indicator distribution



An example of this approach is the normalisation of the '# people in need' indicator, whose raw distribution is shown in Figure 4. The maximum value is 2.6 million. There are also several empty records from early 2022, when people-in-need figures were not computed/available.

The 0.95 and 0.05 quantiles are 1.6 million and 0, respectively. The data is clearly positively skewed. With the default approach (clip the numbers in the 0 to 1.6 million range and apply the logarithmic transformation and min-max normalisation), the distribution would look like Figure 5.

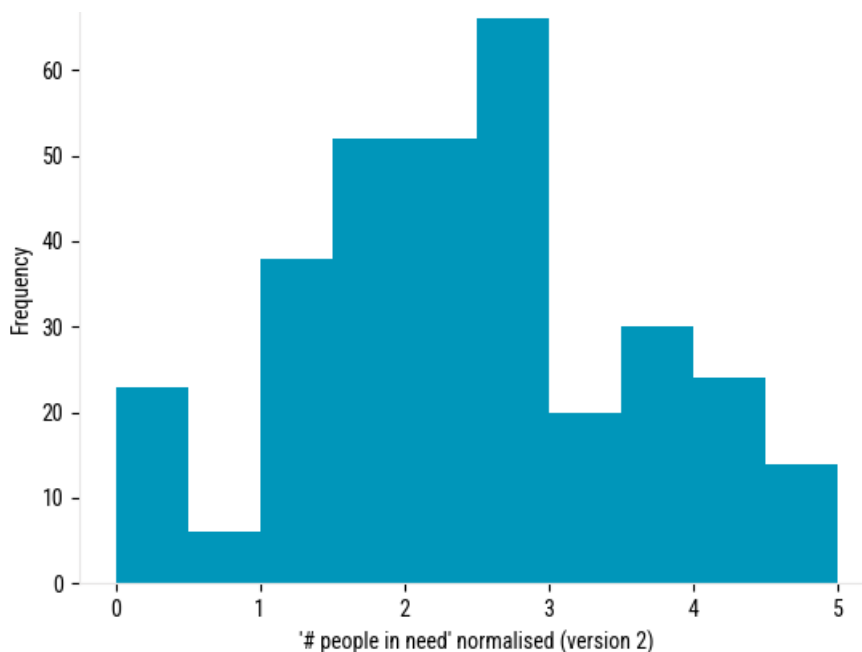
Figure 5. 'People in need' indicator distribution normalised following the approach explained in the previous section



As most of the values were around the 100,000–1,000,000 range, the 0–5 scores gave either very high scores or 0 for cases where there were data gaps (e.g. Crimea region in the initial months of the conflict).

To obtain more evenly distributed values, the minimum value for the clipping was set to 100,000, which corresponded to the 0.07 quantile, and the maximum was set to the actual maximum of the distribution. The final distribution is shown in Figure 6.

**Figure 6. 'People in need' indicator distribution normalised following a case-specific approach**



### Normalisation relative to other indicators

Some of the variables were derived by dividing the absolute value by another indicator, typically the corresponding overall amount of land/total population. This was done to put the number into context, as the absolute value per se did not fully cover the relative situation per oblast and could contribute to smaller scores for smaller oblasts with lower amounts of land and/or population.

Hence, these relative variables were averaged with their absolute counterparts to complement the absolute numbers of people/land affected with the relative numbers to form a more comprehensive indicator.

This process can be illustrated considering the 'I1.4 - # km<sup>2</sup> controlled by Russian forces' indicator from 'D1 - conflict development'.

This indicator is a combination of these subindicators:

I14.1 - # km<sup>2</sup> controlled by Russian forces

I14.2 - % km<sup>2</sup> controlled by Russian forces.

I14.1 follows the approach described in the subsection above – i.e. the corresponding variable '# km<sup>2</sup> controlled by Russian forces' is normalised, with the minimum and maximum respectively set to the .05 and .95 quantiles.

I14.2 is computed by dividing '# km<sup>2</sup> controlled by Russian forces' by '# km<sup>2</sup> landmass affected' and normalising it to obtain a value in the 0–5 range.

The score for I1.4 is obtained as the arithmetic average of I14.1 and I14.2.

The derived variables are not present in the ACAPS Ukraine Master Dataset but computed from other variables contained there. Table 2 shows the derived variables in the model.



Table 3. List of derived variables

VARIABLE NAME	FORMULA	SOURCES	NOTES
% km <sup>2</sup> controlled by Russian forces	$\frac{\{\#km2\ controlled\ by\ russian\ forces\}}{\{\#km2\ landmass\ affected\}}$	ISWN; Live UA Map; ISW; OCHA	
% km <sup>2</sup> unconfirmed control	$\frac{\#km2\ unconfirmed\ control}{km2\ landmass\ affected}$	ISWN; Live UA map; ISW; OCHA	
% registered IDPs	$\frac{\#registered\ IDPs}{\#total\ population}$	IOM; SSS; UNFPA; UNHCR	
% people affected	$\frac{\#people\ affected}{\#total\ population}$	OCHA; SSS; UNFPA; UNHCR; IOM	
% people in N° humanitarian condition	$\frac{\#people\ in\ N^{\circ}\ humanitarian\ condition}{\#total\ population}$	OCHA; SSS; UNFPA; UNHCR; IOM	By N°, we mean 1–5 hum. conditions

## Aggregation

The aggregation method was based on the pillar structure and was identified through the input received during the May hackathon and feedback from contextual experts.

Dimensions were built by grouping together variables that were conceptually similar and/or that belonged to the same topic. To form the dimensions, the normalised variables on the 0–5 scale were arithmetically averaged (with weights in case of D3). The dimensions selected for the framework are as follows.

D1 - conflict development. This dimension groups conflict-related indicators that come from ACLED: fatality rate, battles, explosions, km2 controlled by Russian forces, km2 of unconfirmed control.

D2 - conditions of affected people. This dimension contains information about the distribution of people affected by the severity of conditions. The D2 dimension takes as input the number of people in need in the different severity levels and aggregates them following an approach similar to the one used in the INFORM Severity Index, with the exception of geometric averages, which are in this case substituted by arithmetic averages, and the minima and maxima in the normalisation that are adapted to the specific case of Ukraine oblasts. The specific aggregation method for this dimension is detailed in the D2 - Conditions of Affected People Aggregation section below.

D3 - extent of the crisis. This dimension contains information about the extent of the crisis: number of people affected, number of registered IDPs, number of damage to civilian infrastructure.

D4 - humanitarian access. This dimension has a single indicator, which is itself a composite indicator: the ACAPS Ukraine Subnational Access Score.

D1, D2, and D4 are equal-weight averages of their indicators, while D3 has different weights assigned to each indicator. The four dimensions were then averaged with equal weights to obtain the final score of the ACAPS Severity Model. The schema below shows how the indicators form dimensions and how dimensions form the final index score.

Table 4. Aggregation schema of dimensions

D1 - CONFLICT DEVELOPMENT (0.25)	D2 - CONDITIONS OF AFFECTED PEOPLE (0.25)	D3 - EXTENT OF THE CRISIS (0.25)	D4 - HUMANITARISN ACCESS (0.25)
<ul style="list-style-type: none"> <li>• Fatality rate</li> <li>• Battles</li> <li>• Explosions</li> <li>• KM<sup>2</sup> controlled by russian forces</li> <li>• KM<sup>2</sup> unconfirmed control</li> </ul>	<ul style="list-style-type: none"> <li>• People in need</li> <li>• Concentration of conditions</li> </ul>	<ul style="list-style-type: none"> <li>• People affected (0.4)</li> <li>• Registered IDPs (0.2)</li> <li>• Total infrastructure damage (0.4)</li> </ul>	<ul style="list-style-type: none"> <li>• ACAPS Ukraine Subnational Access Score</li> </ul>

## D2 - Conditions of affected people aggregation

D2 is an arithmetic average of 'I2.1 - people in need' and 'I2.2 - concentration of conditions'. While I2.1 uses the normalisation techniques described in the Normalisation section (min-max normalisation, log transformation), I2.2 requires additional explanation.

The computation of I2.2 starts from the '% people in level (1–5) humanitarian conditions' derived from absolute numbers divided by total population numbers (in the case of the Ukraine crisis, the whole population is considered exposed).

Variables for levels 1–5 of humanitarian conditions follow the INFORM Severity Index methodology and are explained in Table 1 of the *INFORM Severity Index: Concept and Methodology note*.

After computing for the derived variables (% people in N<sup>o</sup> humanitarian condition), the cumulative relative frequency is computed: for each level, this is the sum of the percentage of the population in that level plus the percentages of the populations in all the levels above. The score for this indicator is the highest level that has a cumulative relative frequency bigger or equal to a threshold, which is set at 5%, as shown in Table 3 of the *INFORM Severity Index: Concept and Methodology note*.

## Weighting

To compensate for different aspects of each indicator/dimension, the reliability of each variable, update periods, and data scarcity in occupied territories were analysed.

All indicators were averaged with equal weights to form the respective dimensions, except for the number of registered IDPs. This was motivated by the fact that the only data available monthly was the number of officially registered IDPs, which is likely to be an underestimate in occupied areas of Ukraine given the lack of the government service coverage and overall complexity of registering as an IDP closer to the front lines. Estimated IDP figures are available in IOM's *Ukraine General Population Survey*, but given that this is updated quarterly only, the choice for the ACAPS Ukraine Master Dataset and for the ACAPS Ukraine Severity Model was to consider the number of registered IDPs.

After studying the data reliability, update frequency, spatial granularity, and statistical properties of the model, all dimensions were set to have equal weights.

While these decisions were motivated by the structure and quality of the data, the robustness and sensitivity analysis (described in a later section) evaluated their impact on the ranking of the regions.

## Final score structure

After aggregating and weighting indicators into dimensions (each normalised in a standard range of 0–5), we averaged the values of each dimension. This average is the final score of the severity model, a number that represents the severity assessment for admin1 regions of Ukraine for a particular month, ranging from very low (0) to very high (5). This categorisation is aligned with the *INFORM Severity Index concept note*.

Table 5. categorisation of the final score

NUMERICAL LEVEL	DESCRIPTIVE LEVEL
5	Very high
4	High
3	Medium
2	Low
1	Very low

## STATISTICAL ANALYSIS

Statistical analysis can be used to explore the relationships present in the data, guide the choices of weights and different aggregation schemes, and verify the impact of these choices. One of the tools used to check the soundness of the framework is the correlation matrix, which gives pairwise Pearson's correlation coefficients between the variables.

The higher the correlation index between the variables is, the more these two variables tend to 'move in the same way' when comparing different oblasts or different months. The subsection below presents the coefficients of the dimensions of the model and of the indicators that compose it. These studies were done using data from 2023.

### Correlation matrix of dimensions

The correlation matrix for the different dimensions of the model is shown in Figure 8. It is possible to observe a high correlation coefficient between D1 and D4 and between D2 and D3.

The correlation between 'D1 - conflict development' and 'D4 - humanitarian access' is expected, as these dimensions share a portion of indicators and, at an intuitive level, the presence of battles, explosions, casualties, and other conflict-related indicators results in tighter humanitarian access constraints.

The correlation case between D2 and D3 is less obvious: D2 measures the distribution and severity of people's conditions, and D3 contains the number of people affected, registered IDPs, and total infrastructure damage. The investigation of the correlation between the indicators presented in the next paragraph shows that two indicators that are particularly correlated are the number of people in need and the number of people affected, which belong to D2 and D3, respectively.

**Table 6. Correlation matrix of dimensions**

INDICATOR	D1 - CONFLICT DEVELOPMENT	D2 - CONDITIONS OF AFFECTED PEOPLE	D3 - EXTENT OF THE CRISIS	D4 - HUMANITARIAN ACCESS	SEVERITY
D1 - conflict development	1	0.515806	0.551428	0.925512	0.839754
D2 - conditions of affected people	0.515806	1	0.862225	0.675496	0.866717
D3 - extent of the crisis	0.551428	0.862225	1	0.703735	0.87049
D4 - humanitarian access	0.925512	0.675496	0.703735	1	0.930928
Severity	0.839754	0.866717	0.87049	0.930928	1

### Correlation matrix of indicators

To ensure that the current aggregation makes sense, correlation coefficients between the indicators of each dimension were studied. From Figure 9, we can observe the following.

Indicators under 'D1 – conflict development' showed mostly high correlation figures (+0.54 to +0.916 range).

Indicators under 'D2 - conditions of affected people' showed a 0.7 Pearson correlation.

Under 'D3 - impact of the crisis', the situation was different as the range was +0.16 to +0.71. The least correlated indicator was 'I3.3 - total infrastructure damage'.

From the correlation matrix that we have, it is clear that variables under D1 and D2 were highly correlated. This supports aggregating them together. Variables under D3 had the least amount of correlation – specifically, 'I3.3 - total infrastructure damage' had a 0.16 correlation coefficient with 'I3.2 - registered IDPs'. This could be explained by the logic involved in the movement of people, as most people tend to move and register in safer regions of Ukraine, so we could expect the highest number of registered IDPs in regions where there is less infrastructure damage.

Notably, there was a negative correlation (-0.30) between 'I3.2 - registered IDPs' and 'I1.4 - km2 controlled by Russian forces'. From the data analysed, a high score for I1.4 tended to result in a low I3.2 score, and vice versa. This is also expected, as people tend to move from occupied regions to safer ones that are not controlled by Russian forces. The model structure includes different dimensions exactly to cover different aspects of severity. One region could be facing a severe situation because of battles and occupied areas, while the situation in another region could be severe in terms of the IDP influx.

The correlation between 'I2.1 - people in need' and 'I3.1 - people affected' is also expected, as the people-in-need indicator is a subset of people affected by definition in the INFORM severity methodology.

Table 7. correlation matrix of indicators

INDICATOR	I1.1 - FATALITY RATE	I1.2 - BATTLES	I1.3 - EXPLOSIONS	I1.4 - KM <sup>2</sup> CONTROLLED BY RUSSIAN FORCES	I1.5 - KM <sup>2</sup> UNCONFIRMED CONTROL	I2.1 - PEOPLE IN NEED	I2.2 - CONCENTRATION OF CONDITIONS	I3.1 - PEOPLE AFFECTED	I3.2 - REGISTERED IDPS	I3.3 - TOTAL INFRASTRUCTURE DAMAGE	SEVERITY
I1.1 - fatality rate	1	0.916671	0.817915	0.735085	0.891861	0.457785	0.577175	0.460739	0.143619	0.633046	0.848376
I1.2 - battles	0.916671	1	0.721428	0.700429	0.895886	0.513344	0.618119	0.484922	0.225264	0.490352	0.845955
I1.3 - explosions	0.817915	0.721428	1	0.547852	0.717897	0.500595	0.577056	0.573426	0.177887	0.783698	0.85104
I1.4 - km <sup>2</sup> controlled by Russian forces	0.735085	0.700429	0.547852	1	0.735642	-0.03221	0.099911	-0.03541	-0.30895	0.268463	0.427136
I1.5 - km <sup>2</sup> unconfirmed control	0.891861	0.895886	0.717897	0.735642	1	0.365854	0.644916	0.41139	0.059568	0.540824	0.82306
I2.1 - people in need	0.457785	0.513344	0.500595	-0.03221	0.365854	1	0.690525	0.912959	0.712883	0.488119	0.74346
I2.2 - concentration of conditions	0.577175	0.618119	0.577056	0.099911	0.644916	0.690525	1	0.725973	0.601296	0.536167	0.851714
I3.1 - people affected	0.460739	0.484922	0.573426	-0.03541	0.41139	0.912959	0.725973	1	0.714027	0.50415	0.779005
I3.2 - registered idps	0.143619	0.225264	0.177887	-0.30895	0.059568	0.712883	0.601296	0.714027	1	0.162034	0.473484
I3.3 - total infrastructure damage	0.633046	0.490352	0.783698	0.268463	0.540824	0.488119	0.536167	0.50415	0.162034	1	0.739692
Severity	0.848376	0.845955	0.85104	0.427136	0.82306	0.74346	0.851714	0.779005	0.473484	0.739692	1

## Robustness and sensitivity analysis

While all of the decisions that led to the final structure of the index were based on a study of the available data and discussed with context experts, each of them maintained a degree of arbitrariness. To evaluate the impact of these 'arbitrary' choices on the final score of the model and on the ranking of the oblasts, a sensitivity and robustness analysis was designed.

The essential principle of this analysis was to randomise the parameters that we assigned for normalisation, log transformation, and weights. Repeating this operation thousands of times allowed us to obtain, for each oblast and for a given month, a distribution of scores and a distribution of rankings. The width of these distributions allowed us to estimate the impact of the model choices made.

A random generation of parameters was performed 1,000 times, modifying the following parameters.

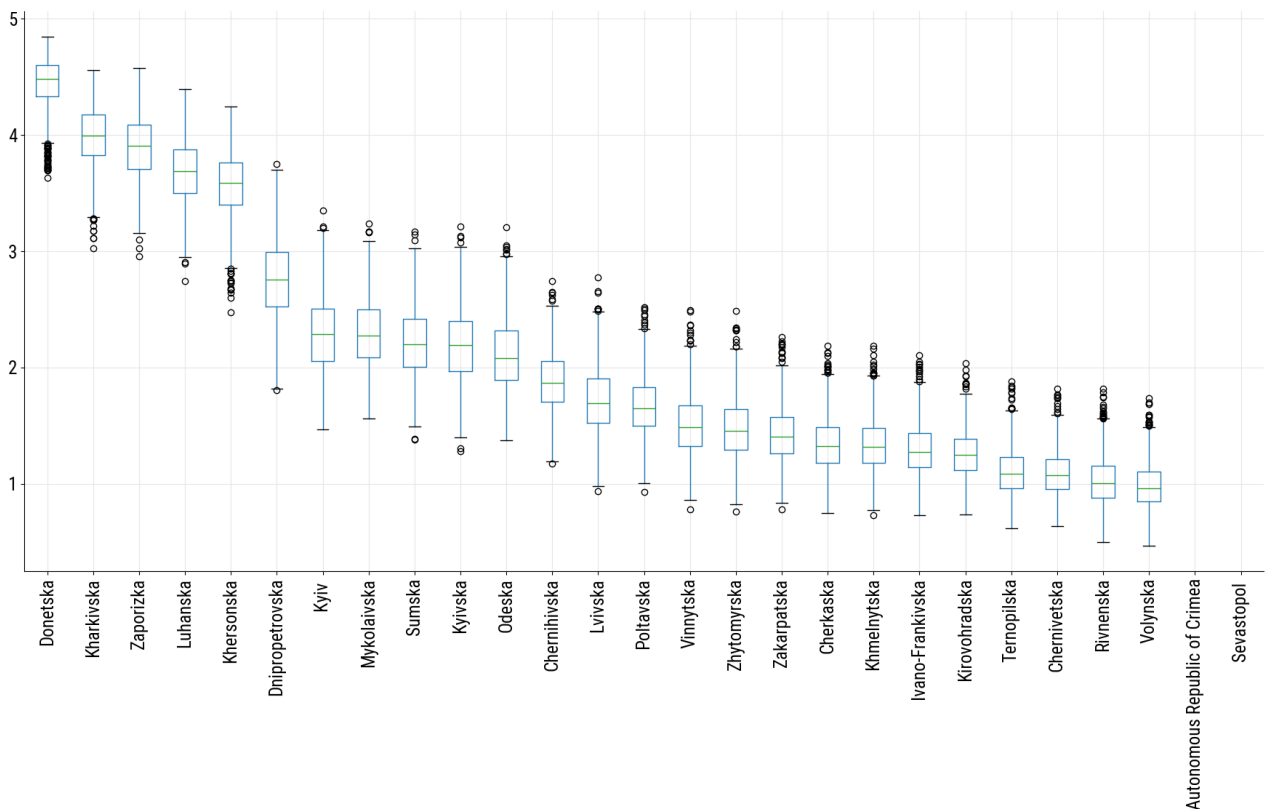
Min and max parameters were randomly selected around the range of the variable's 0.05 and 0.95 quantiles, respectively. The range around the quantiles was selected on a case-by-case basis, as distributions for the different indicators varied.

Log10 transformations were randomly applied.

The weights were randomly set to a value in the (0.1–0.4) range. The weights did not have to sum up to 1, as the value for the weighted average was computed using the formula:

The model score for December 2023 was calculated for each iteration. Figure 10 shows a box plot with the minimum, maximum, median, and 0.25 and 0.75 quantiles of the distribution of the scores for each oblast.

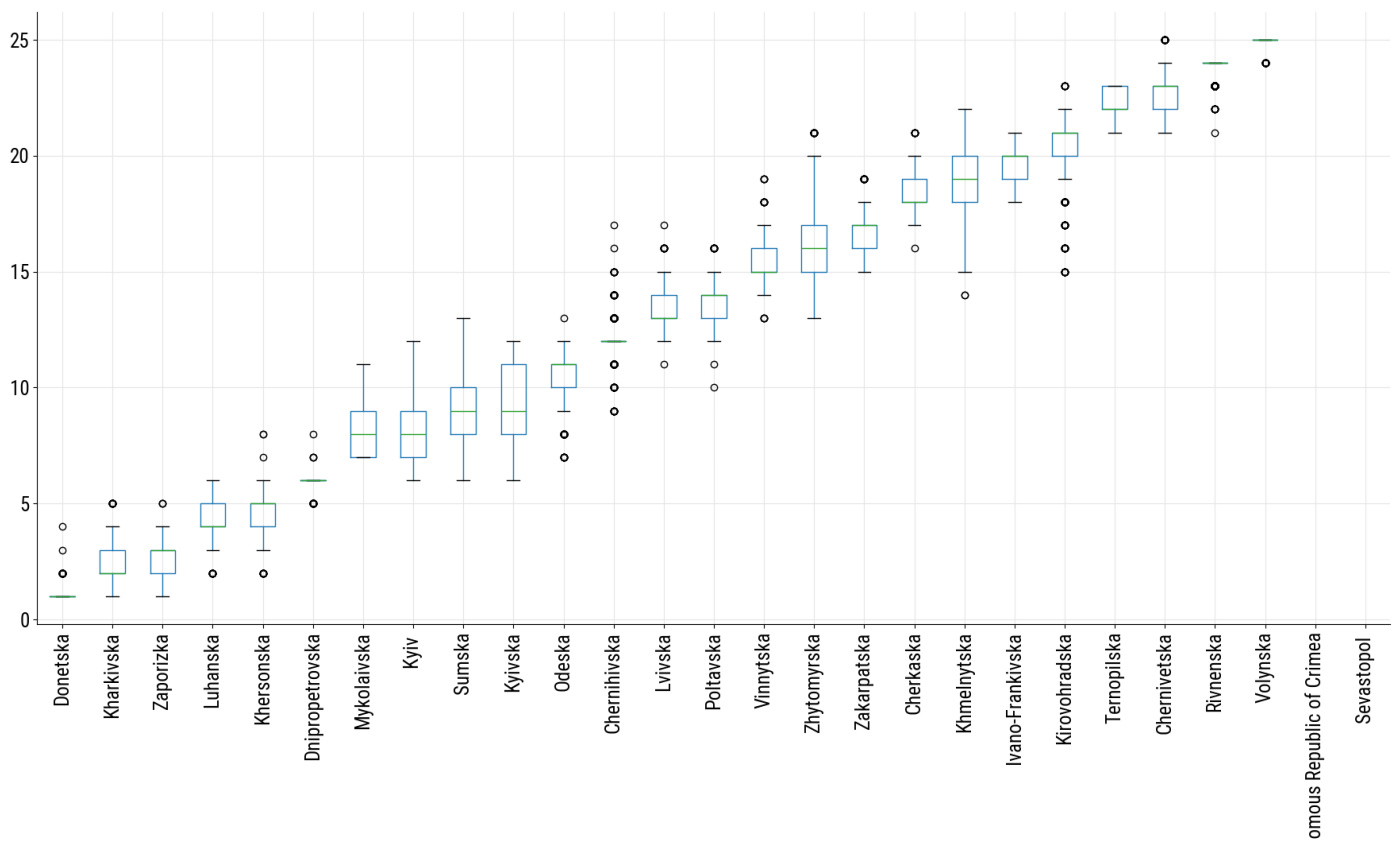
Figure 7. Distribution of severity



We can see that the distribution the scores obtained for each oblast with the different random choices is quite narrow. This indicates that the final score is stable regardless of the individual choices that we make when normalizing and aggregating the different indicators.

Figure 11 shows the distribution of the rankings.

Figure 8. Distribution of severity rank



As expected, the rankings showed an even narrower distribution than scores. This is proof of the stability of the model, regardless of the choices applied.

It is also possible to study the ranking for individual dimensions, shown in the figures below.

Figure 9. Dimension 1 ranking

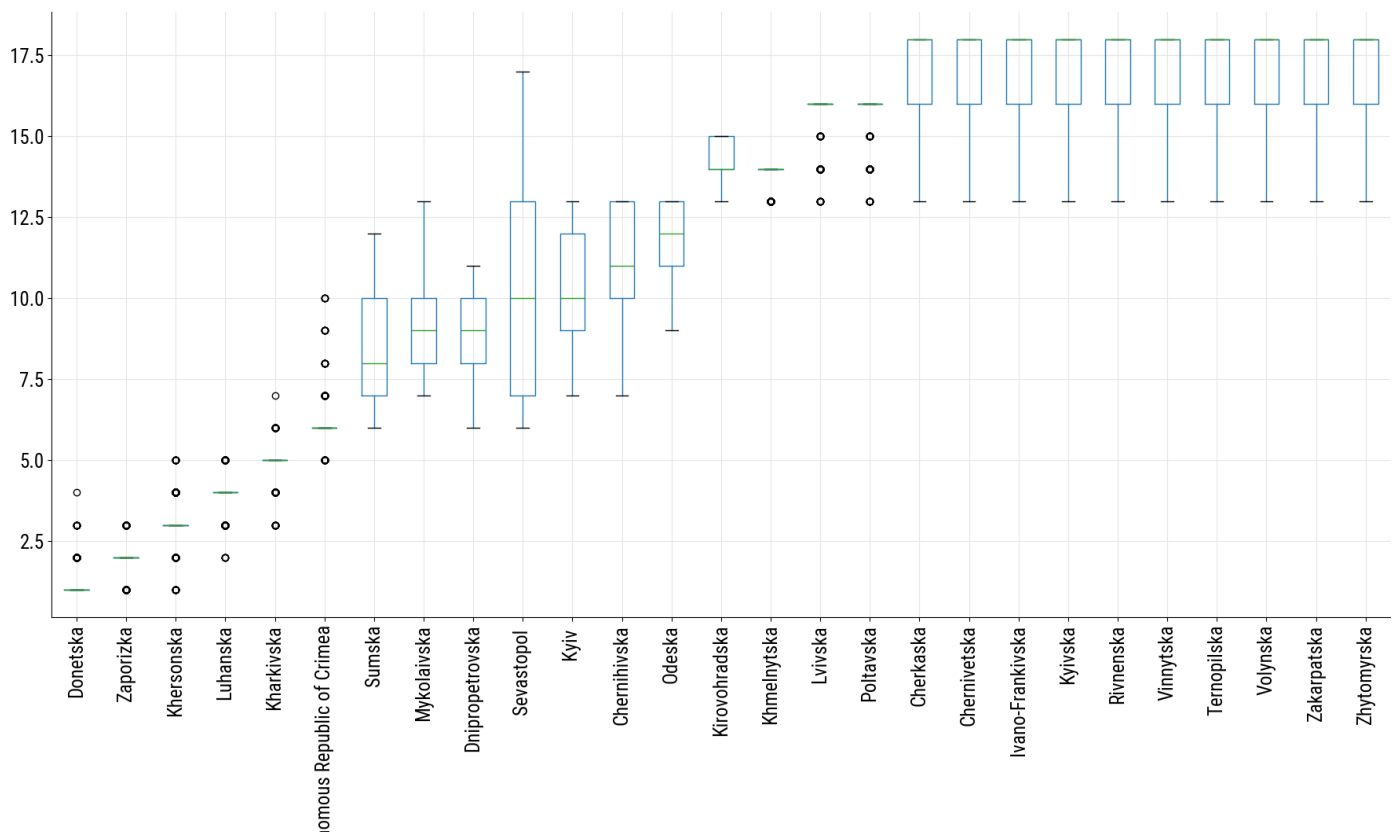


Figure 10. Dimension 2 ranking

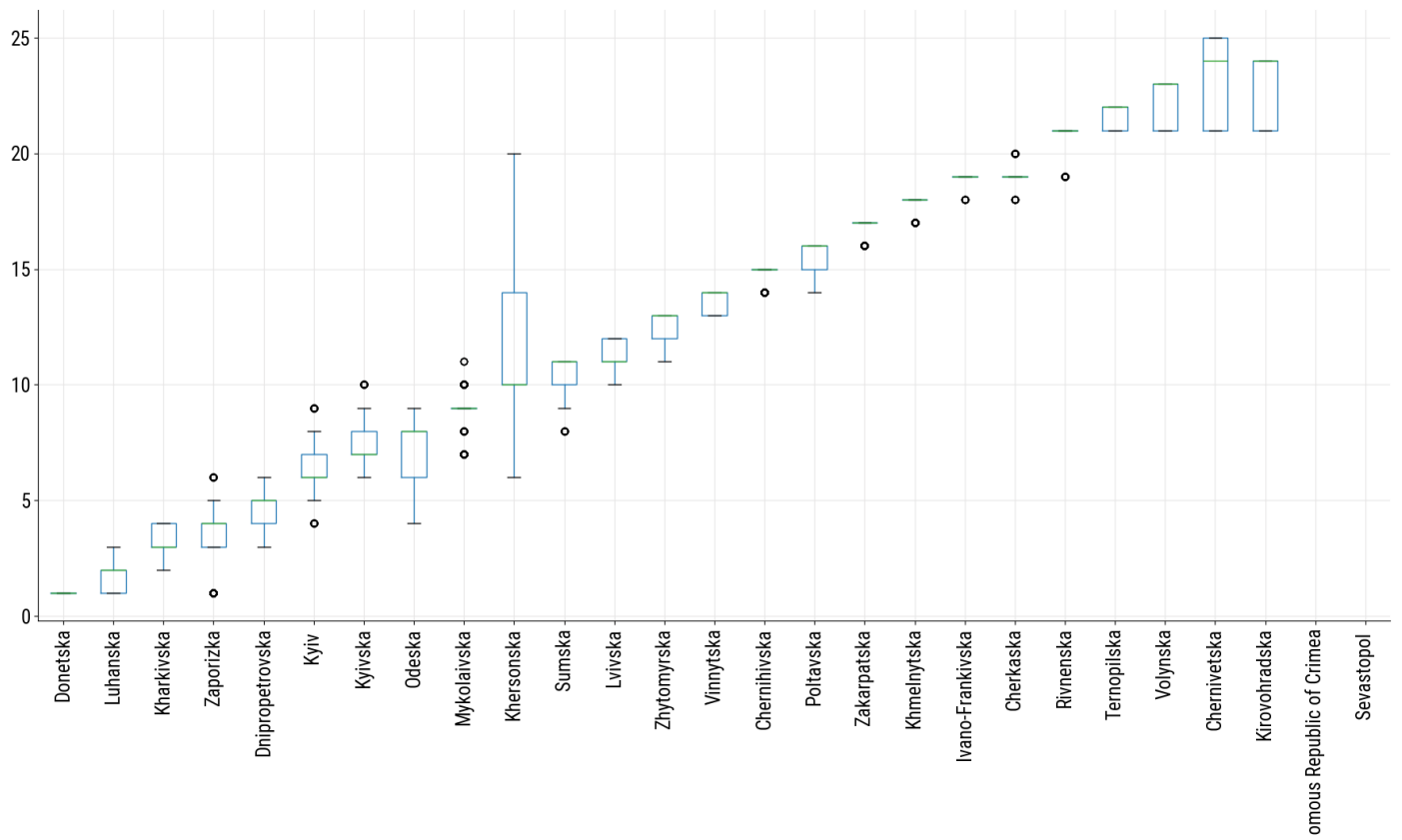
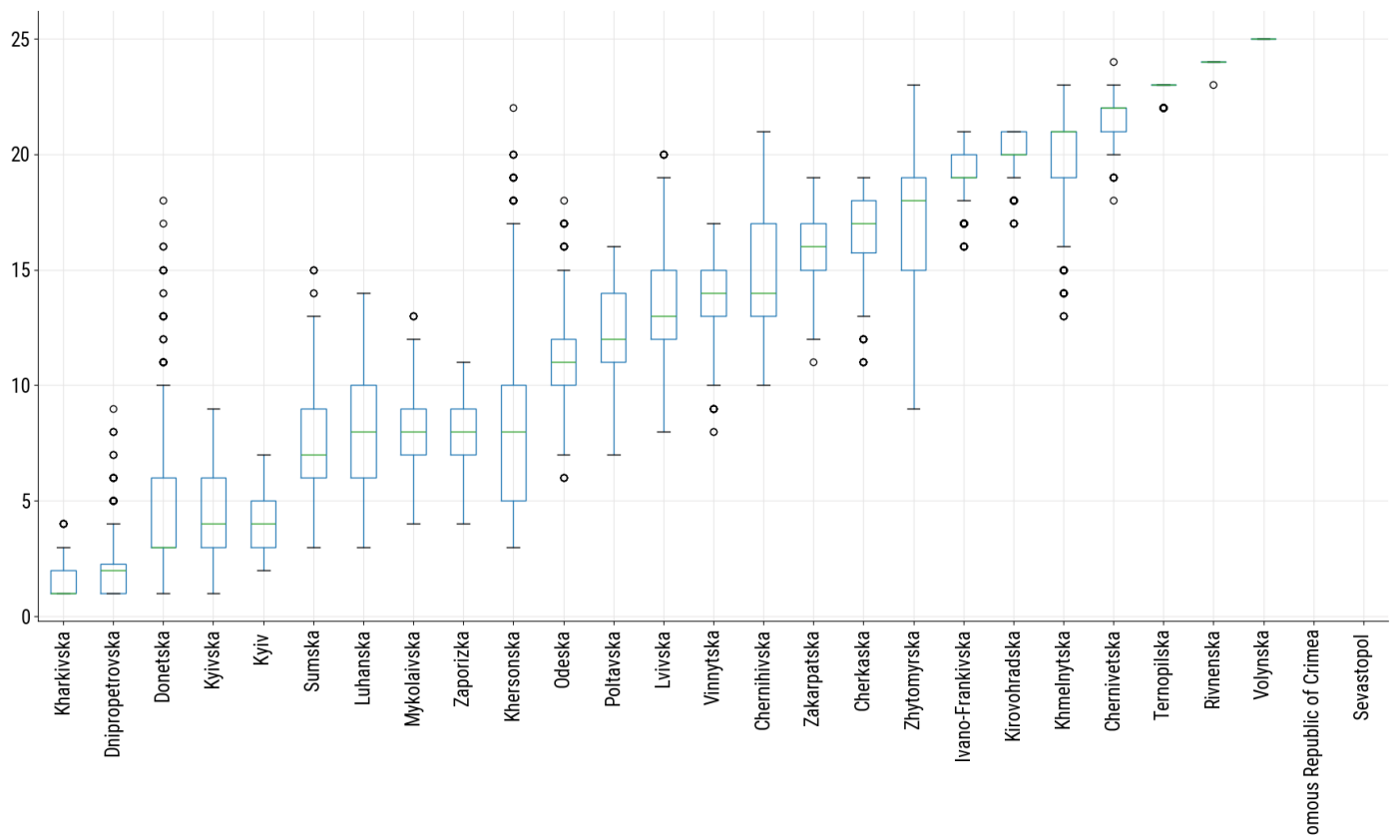


Figure 11. Dimension 3 ranking



## LIMITATIONS

### Repeated indicators

Some of the indicators from dimensions other than D4 were partially used (even if in modified forms) in the ACAPS Ukraine Subnational Access Score, leading to a 'double counting' of these indicators.

This choice was done consciously to create indicators and dimensions highlighting aspects particularly relevant for the humanitarian crisis in Ukraine (in this case, conflict is the main driver of the crisis). The ACAPS Ukraine Subnational Access Score is a composite indicator whose methodology is available on the ACAPS website and is based on the ACAPS Humanitarian Access methodology. Conflict-related indicators are part of this model, as active conflict has an impact on humanitarian access. That said, these indicators are averaged with many others, and their weight is reduced in the final score. In a conflict-driven humanitarian crisis, such as the one in Ukraine, these indicators assume a much larger importance in the final determination of the severity.

From the point of view of the final score of the model, this choice is in practice equivalent to assigning a higher weight to the indicators that appear both as part of the ACAPS Ukraine Subnational Access model and as standalone indicators in the ACAPS Ukraine Severity Model. This effect is small given that, as mentioned previously, the ACAPS Ukraine Subnational Access model contains many more indicators that are all used to obtain the final score.

The full list of indicators that are present both as standalone indicators in the ACAPS Ukraine Severity Model as well as in the ACAPS Ukraine Subnational Access score is presented below.

**Table 4. List of indicators with some level of repetition**

INDICATOR'S NAME	DIMENSIONS	NOTES
# total infrastructure damage	D3, D4	# total infrastructure damage, which is an indicator of D3, is partially present in the ACAPS Ukraine Subnational Access for the computation of subindicators 7.2 and 9.2. The number of infrastructure damage is used for two corresponding indicators using this type of event:  I7 - insecurity or hostilities affecting humanitarian assistance  I9 - physical constraints in the environment
# battles	D1, D4	Access score computation uses the total number of ACLED events for the calculation of S7.1. The # battles variable corresponds to the count of a subset of the total number of ACLED events.
# explosions	D1, D4	Access score computation uses the total number of ACLED events for the calculation of S7.1. The # explosions variable corresponds to the count of a subset of the total number of ACLED events.

### Data limitations: frequency of update

As the model aims to provide monthly severity assessments for each region of Ukraine, time granularity is one of the most important factors of variables being included/excluded. The minimum threshold of inclusion is the quarterly update of data sources, but there are exceptions for the variables crucial to the humanitarian response and for which no valid alternatives are available.

In particular, this applies to the number of people in need, people affected, and people in 1–5 humanitarian conditions. These are taken from the Humanitarian Needs Overview or Humanitarian Needs and Response Plan reports that are updated yearly. The exception also applies to total population figures, also updated yearly.

### Other data limitations

Some indicators present in the ACAPS Ukraine Master Dataset, in particular the socioeconomic variables (e.g. basket prices, inflation rates), did not meet the needed granularity criteria. Most of the data is either updated only yearly or is only available at the national level. The only socioeconomic indicator in the ACAPS Ukraine Master Dataset that meets the granularity criteria required for the ACAPS Ukraine Severity Model is the number of people unemployed. This was also not included in the model, however, because, by itself, it does not fully represent the socioeconomic situation. At the same time, it is potentially underestimated in Ukraine, especially in the Russian-occupied territories, as not all unemployed people register as such or because of the high complexity of registering in occupied/frontline territories.



Another data limitation that has an impact on the framework relates to the '# registered IDPs' indicator. While in most cases the model uses equal-weight averages, this indicator is associated with a weight lower than other indicators' in the same dimension. This is because of the complex situation around access restrictions and information gaps on the number of registered IDPs per oblast, which make this indicator less reliable in regions with limited government service coverage/availability. Decreasing the weight of this indicator allows for the use of the information while limiting its impact on the final score.

Notably, data coverage in Ukraine is limited in the Russian occupied areas, where final scores might be underestimated. Given the huge data gaps, the scores for Crimea and Sevastopol were not provided, as these regions lack updates on some of the key variables used in the model.

## ANNEX I. HOW TO USE THE INDEX: A HANDS-ON GUIDE

---

To illustrate how the index can be used, this annex contains code snippets that show you how to retrieve the data, analyse it in Python, and visualise the results.

This is done using the ACAPS API in the Jupyter Notebook workspace. To follow this guide, please register at [acaps.org](https://acaps.org), have Python 3 and Jupyter Notebook installed on your computer, and install the required Python packages (Requests, pandas, GeoPandas, and Ukraine's admin1 shapefiles for the map section).

The curl commands can be run either directly in the bash/command prompt (that has to support curl, meaning Windows 10 build 17063 or higher) or, for convenience, using the same options in the Jupyter Notebook with the '!' character inclusion at the start (!curl -X etc.).

### Getting the data

1. Firstly, the code starts with importing the needed libraries for Python 3. They should be installed for your environment beforehand:

```
import requests #library for loading data from the ACAPS API
import pandas as pd #library for working with tabular data
import matplotlib.pyplot as plt #library for visualisations
import geopandas as gpd #library for working with geographic data
```

2. Then, you need to have an authorisation token to request the Ukraine Severity Model results. For it, you need to run a simple curl command with your credentials:

```
curl -X POST "https://api.acaps.org/api/v1/token-auth/" -d
"username=test%40example.com&password=xxxxxxxxxxx"
```

More on how to use the API could be found in the ACAPS API documentation.

3. After saving it in a dictionary variable, for example:

```
headers = {'Authorization': 'Token your_token_here'}
```

4. You need to save the URL you are interested in (ukraine/severity-model/Nov2023, for example):

```
url = 'https://api.acaps.org/api/v1/ukraine/severity-model/Nov2023/'
```

5. You are ready to retrieve the data from the ACAPS API. First, you need to run get method from the Requests library using the URL and header variables that we defined above:

```
r = requests.get (url, headers=headers)
```

6. Then, we need to get the results from the dictionary into a dataframe. The results key contains the exact Ukraine Severity Model info:

```
df_sev_index = pd.DataFrame.from_dict(r.json()['results'])
```

So, we are ready to work with a dataframe object that has Nov2023 data.

### Transforming the dataframe

After getting the data into a dataframe, we could filter for the month that we are interested in (Nov2023 in our example) and order the results by Severity rank:

```
df_sev_index.loc[(df_sev_index['Month']=='Nov2023') & (df_sev_index['Indicator']=='Severity'), ['Oblast', 'Month', 'Rank', 'Indicator', 'Figure']].sort_values('Rank')
```

The results then are returned as:

	OBLAST	MONTH	RANK	INDICATOR	FIGURE
4988	Donetska	Nov2023	1.0	Severity	4.545
4986	Kharkivska	Nov2023	2.0	Severity	3.9425
4987	Khersonska	Nov2023	3.0	Severity	3.8525
4970	Zaporizka	Nov2023	4.0	Severity	3.81
4981	Luhanska	Nov2023	5.0	Severity	3.55
4991	Dnipropetrovska	Nov2023	6.0	Severity	2.68
4972	Sumska	Nov2023	7.0	Severity	2.2025
4979	Mykolaivska	Nov2023	8.0	Severity	2.1825
4992	Chernihivska	Nov2023	9.0	Severity	1.9125
4976	Odeska	Nov2023	10.0	Severity	1.9075
4983	Kyivska	Nov2023	11.0	Severity	1.8775
4982	Kyiv	Nov2023	12.0	Severity	1.82
4978	Poltavska	Nov2023	13.0	Severity	1.46
4967	Zhytomyrska	Nov2023	14.0	Severity	1.45
4985	Khmelnyska	Nov2023	15.0	Severity	1.44
4973	Vinnytska	Nov2023	16.0	Severity	1.4
4994	Cherkaska	Nov2023	17.0	Severity	1.3775
4980	Lvivska	Nov2023	18.0	Severity	1.37
4969	Zakarpatska	Nov2023	19.0	Severity	1.245
4989	Ivano-Frankivska	Nov2023	20.0	Severity	1.2425
4984	Kirovohradska	Nov2023	21.0	Severity	1.2325
4990	Chernivetska	Nov2023	22.0	Severity	1.11
4974	Ternopilka	Nov2023	23.0	Severity	1.095
4977	Rivnenska	Nov2023	24.0	Severity	1.0775
4971	Volynska	Nov2023	25.0	Severity	1
4975	Sevastopol	Nov2023	26.5	Severity	NaN
4993	Autonomous Republic of Crimea	Nov2023	26.5	Severity	NaN

As you can see above, we have Oblast, Month, Rank (of the corresponding variable, Severity in this case), and variable value.

This result can then be saved as an .xlsx or .csv by adding an extra line:

```
df_sev_index_rank.to_excel('sev_ranks_nov_23.xlsx')
```

That said, tables are not always the best choice for exploring or analysing data, so we also share a few examples of visualisations with chronological Severity score for different regions of Ukraine.

First, you need to choose the oblasts that we are interested in; south regions are chosen as an example.

To save south regions' names in a list variable:

```
south_regions = ['Khersonska', 'Mykolaivska', 'Odeska']
```

Choosing months for the visualisation, keep in mind that the xticks variable should be equal to the length of periods\_full variable:

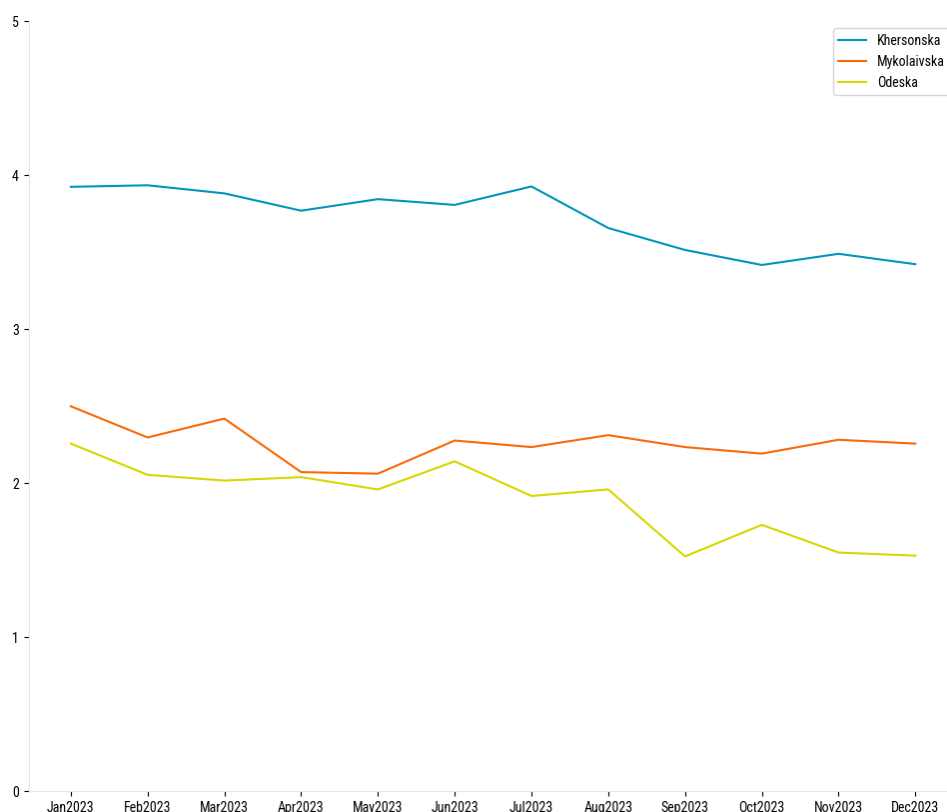
```
xticks=range(12)
```

```
periods_full = ['Jan2023', 'Feb2023', 'Mar2023', 'Apr2023', 'May2023',  
               'Jun2023', 'Jul2023', 'Aug2023', 'Sep2023', 'Oct2023',  
               'Nov2023', 'Dec2023']
```

Then we filter through the working dataframe to show only relevant information:

```
for region in south_regions:  
    access = df_sev_index.loc[(df_sev_index['Oblast']==region)  
    & (df_sev_index['Indicator']=="Severity"), 'Figure'].plot(legend=True,  
    linestyle='solid', figsize=(12, 10), markersize = 12, xticks=xticks, use_index=False,  
    ylim=(0.0, 5.0))  
    access.set_xticklabels(periods_full)  
    plt.legend(south_regions, loc=1)
```

Finally, the results:



To show the data on the map, you have to install GeoPandas, as well as the Ukraine shapefile to visualise the results. We suggest OCHA's boundaries for getting the Ukraine adm1 shapefile.

First, you need to extract the files from the archive, and then save the path to the adm1 shapefile in the ADM1\_PATH variable.

```
ADM1_PATH = r'/your_working_folder/borders/ukr_admbnda_adm1_sspe_20221005.shp'
```

Then, you need to read it using the `gpd.read_file()` command.

```
borders_adm1 = gpd.read_file(ADM1_PATH, SHAPE_RESTORE_SHX = 'YES')
```

The last command is to check that everything has gone correctly:

```
borders_adm1['ADM1_EN'].unique()
```

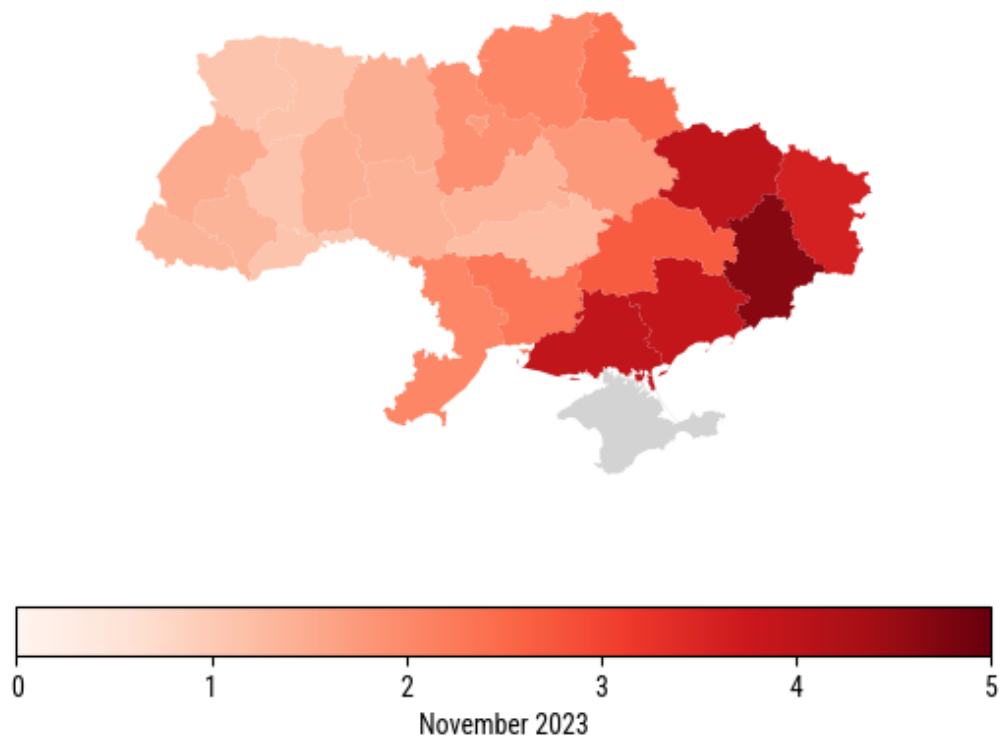
Lastly, we need to merge the geographical GeoPandas dataframe with ours:

```
df_map = pd.merge(borders_adm1, df_sev_index, how = 'right', left_on='ADM1_EN',  
right_on='Oblast')
```

After connecting the GeoPandas frame with our working dataframe, we are ready to filter for the relevant data and show it on the map:

```
df_map[df_map['Month']=='Nov2023'].plot("Severity", legend = True, cmap='Reds', vmin = 0.0,  
vmax = 5.0, missing_kwds={'color': 'lightgrey'})
```

The code below maps the severity scores for Nov2023 using shades of red. The results are shown below:



The Autonomous Republic of Crimea is in grey as the final Severity score is not available given current data gaps.

To save the image, you need to add additional lines:

```
graph = df_map[df_map['Month']=='Nov2023'].plot("Severity", legend = True, cmap='Reds', vmin  
= 0.0, vmax = 5.0)  
  
fig = graph.get_figure()  
  
fig.savefig('path_where_file_is_saved')
```

## ANNEX II. TRANSFORMATIONS

Table 8 summarises all the data transformations applied in the model.

**Table 8. Transformations of the variables**

VARIABLE	MIN-MAX APPLIED	WEIGHT APPLIED	LOG TRANSFORMATION	ADDITIONAL CHANGES
D1 - conflict development	No	0.25	No	
D2 - conditions of affected people	No	0.25	No	
D3 - extent of the crisis	No	0.25	No	
D4 - humanitarian access	No	0.25	No	
I1.1 - fatality rate	Yes (1, 750)	No weight applied (same as equal weights)	Yes	
I1.2 - battles	Yes (1, 150)	No weight applied	Yes	
I1.3 - explosions	Yes (1, 550)	No weight applied	Yes	
I1.4 - km <sup>2</sup> controlled by Russian forces	No	No weight applied	No	Arithmetic average of I1.4.1 and I1.4.2
I1.4.1 - # km <sup>2</sup> controlled by Russian forces	Yes (1, 25,000)	No weight applied	Yes	
I1.4.2 - % km <sup>2</sup> controlled by Russian forces	Yes (0, 1)	No weight applied	No	The variable is calculated by dividing I1.4.1 by the total landmass affected
I1.5 - km <sup>2</sup> unconfirmed control	No	No weight applied	No	Arithmetic average of I1.5.1 and I1.5.2
I1.5.1 - # km <sup>2</sup> unconfirmed control	Yes (1, 270)	No weight applied	Yes	
I1.5.2 - % km <sup>2</sup> unconfirmed control	Yes (0, 0.01)	No weight applied	No	The variable is calculated by dividing '# km <sup>2</sup> unconfirmed control' by '# km <sup>2</sup> landmass affected'
I2.1 - people in need	Yes (100,000, 2,500,000)	No weight applied	Yes	
I2.2 - concentration of conditions	-	-	-	The calculation process is explained in the Aggregation section
I3.1 - people affected	-	Yes, 0.4	-	Arithmetic average of I3.1.1 and I3.1.2
I3.1.1 - # people affected	Yes (100,000, 2,600,000)	No weight applied	Yes	
I3.1.2 - % people affected	Yes (0.0, 0.9)	No weight applied	No	
I3.2 - registered IDPs	No	Yes, 0.2	No	Arithmetic average of I3.2.1 and I3.2.2
I3.2.1 - # registered IDPs	Yes (4,000, 500,000)	No weight applied	Yes	
I3.2.2 - % registered IDPs	Yes (0.001, 0.2)	No weight applied	No	
I3.3 - total infrastructure damage	Yes (1, 100)	Yes, 0.4	Yes	

## SOURCES

Nardo, Michela, Michaela Saisana, Andrea Saltelli, Stefano Tarantola, Anders Hoffman, Enrico Giovannini. 2005. Handbook on Constructing Composite Indicators: Methodology and User Guide

Poljansek, K., P. Disperati, L. Vernaccini, A. Nika, S. Marzi, and A. H. Essenfelder. 2020. INFORM Severity Index, EUR 30400 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-23014-4, doi:10.2760/94802, JRC122162.

ACAPS. n.d. "API documentation." Accessed February 28, 2024. [api.acaps.org/using-the-api/](https://api.acaps.org/using-the-api/).

OCHA. n.d. "Ukraine - Subnational Administrative Boundaries." Accessed February 28, 2024. <https://data.humdata.org/dataset/cod-ab-ukr>.