



Khan Ahmad from Ghor province sits in his mud house in Police Rah informal settlement near Herat city. He fled his Taliban-controlled village three years ago for security reasons and now finds himself stranded without money or work. Photo: © UNHCR/Jim Huylebroek, June 2016

PART 3

INSIDE THE GRID

Overcoming data shortfalls

A number of global and regional policy processes have recognised the importance of collecting and analysing credible and transparent data on internal displacement. Such an evidence base is essential as a yardstick against which to measure progress toward implementation of the 2030 Agenda for Sustainable Development, the Agenda for Humanity, the Sendai Framework for Disaster Risk Reduction 2015-2030, the UNFCCC Paris Agreement and other commitments on climate change, the Nansen Initiative's protection agenda for people displaced across borders by disasters, the Valletta Summit action plan and the New Urban Agenda.

There is also a growing demand for evidence to inform the two-year negotiations on the global compacts for safe, orderly and regular migration, and on sharing responsibility for refugees in 2018.

In his report to the World Humanitarian Summit, the previous UN secretary-general stated that:

*Data and joint analysis must become the bedrock of our action. Data and analysis are the starting point for moving from a supply-driven approach to one informed by the greatest risks and the needs of the most vulnerable. National Governments and subregional, regional and international actors need to dedicate significant financial and human resource capacity towards collecting data and monitoring and analysing risk before, during and after crises, particularly in the most risk-prone countries and areas.*²⁶⁵

The UN Statistical Commission (UNSC) has also recognised the need for better data on IDPs. At its 47th session in March 2016, it established an expert group and called for a technical report on official statistics for IDPs and refugees to be prepared in time for its 49th session in 2018.²⁶⁶ Importantly, it has also recognised the need for such data to be more interoperable and account

for the times when IDPs cross international borders to seek protection.

Comprehensive stock and flow data is also needed to monitor progress towards the UN secretary-general's ambitious target of "reducing new and protracted internal displacement by 2030" by at least 50 per cent in ways that "always guarantee voluntariness, dignity and safety" and "never compromise the right to flee."²⁶⁷

IDMC's global data will serve as the baseline against which progress toward this target is measured and to direct attention where it is most needed. In addition to the secretary general, the UN General Assembly and member states have repeatedly underscored the need for global data, and for IDMC to provide it.²⁶⁸

This section of the GRID highlights some of the main challenges IDMC faced this year, and what they mean for future data analysis and global policy monitoring, both in terms of the need to harmonise data collection systems and approaches, and for more investment by governments in monitoring displacement situations over time. A full description of IDMC's accounting is included in the methodological annex.

KEY FINDINGS

Reliable data and analysis are central to the achievement of global and regional development and humanitarian policy processes relevant to IDPs. Demands for systematic data collection, analysis and research have not however been matched by the political will and resources required to meet them. As a result, the current baseline and global picture of internal displacement are currently incomplete.

The time-series data needed to measure progress toward global targets is not collected through to the end of displacement. This means that we do not properly understand how different displacement situations and specific IDPs' vulnerabilities evolve over time. Further gaps include limited geographical scope, exclusion of certain types of displacement, and disaggregation of data by age, sex, location, needs and vulnerabilities.

Without comparable data on different situations and how they have changed over time, there is little evidence to tell us what works. Yet this information is critical to remove the guesswork currently involved in humanitarian and development financing. New and innovative solutions need to be deployed to fill the data gaps and establish a more comprehensive picture of displacement. New "hybrid" approaches that combine event detection with the analysis of time-series data on evolving situations are essential.

Detecting incidents of new displacement needs to be scaled up significantly, employing semi-automated processes that monitor displacement associated with disasters, conflict, violence and development projects. For disasters, more time-series data on people once they have become displaced is key in order to infer both the total number of people displaced by an event and to track the number and needs of displaced people as they evolve over time. In addition, more investment is needed in probabilistic risk modelling for disaster displacement in order to identify and address the drivers of displacement risk. Furthermore, the assessment of displacement risk and the factors that contribute to it should be extended to other contexts.

Internal displacement data

AN INCOMPLETE PICTURE

Despite our best efforts, the GRID does not yet paint a comprehensive picture of internal displacement worldwide (see figure 3.1). This means that our global baseline is still a significant underestimate. Key gaps include the lack of data on all relevant phenomena, our limited ability both to obtain and analyse all of the information that does exist and to systematically identify new incidents of displacement. Without this information, we do not have an accurate measure of how many people have become internally displaced, the reasons they have fled and how long they remain displaced for.

Limits in geographical coverage

We added Algeria, Burkina Faso and Mozambique to our dataset on displacement associated with conflict in 2016, bringing the total number of countries and territories monitored to 56. Our monitoring of displacement associated with disasters covered 118 countries and territories in 2016 and 176 since 2008.

ACCOUNTING TERMS AND CONCEPTS

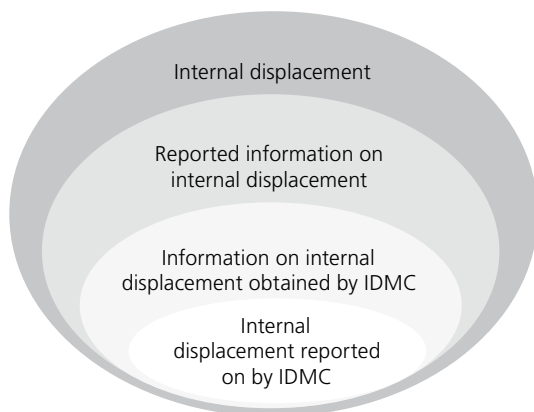
The language we use to describe how we account for internal displacement can seem abstract and far removed from people at the heart of this report. Behind all of our figures are people whose lives have been threatened and disrupted, in many cases severely, by traumatic events. Most displaced people flee their homes and places of residence as a last resort and only in response to life-threatening situations.

Most displacement figures and statistics refer to “stocks” or “flows”.

Concept	Characteristics	Commonly used terms
Stock	The number of people, in this case IDPs, in a given situation and/or location at a particular moment in time	IDPs, refugees, migrants, returnees
Flows	Processes, such as the rate at which people are newly displaced or return over a given period of time	New displacements, returns, resettlements

As a report to UNSC noted: “The production of statistics on [displaced people] requires a clear distinction between stocks and flows.”²⁶⁹ Doing so remains a challenge though, even among national statistics offices and those collecting displacement data. The term “returnees”, for example, can refer to IDPs who are no longer displaced, having returned to their homes or places of origin. It is also used, however, to describe refugees who have returned to their country of origin but who may still be displaced and accounted for as such (see below for more detail).

Figure 3.1: Coverage of IDMC's reporting (not to scale)



We were unable to obtain verifiable quantitative estimates of new displacement or the total stock of IDPs for some countries, despite significant contextual evidence that internal displacement had occurred. This was the case for people displaced by gangs and criminal violence in Brazil and South Africa.

For other countries, we lack the disaggregated data needed to comprehensively account for all incidents of displacement. The Chinese government, for example, publishes its own annual estimates of the number of people evacuated and otherwise displaced by disasters. Their figures have been higher than our own estimates each year, because their aggregate includes disasters about which we have not been able to obtain information. We have been closing this gap every year with increased attention and effort.

Often our geographical coverage is limited for reasons beyond our and our partners' control. Gaps in geographical coverage can occur because insecurity impedes access to IDPs, a challenge our sources face in numerous countries. Sometimes limited resources are to blame. In one country, our source reduced the geographical scope of its data collection in December from 11 to eight provinces for want of funding. In other countries politics play a key role, such as when the government only authorises data collection in certain locations, which are not necessarily where the majority of IDPs are sheltering.

Not all causes of internal displacement are included in the global baseline

The Guiding Principles on Internal Displacement recognise that people may become displaced for a number of reasons including, but not limited to "armed conflict, situations of generalized violence, violations of human rights or natural or human-made disasters."²⁷⁰ Current monitoring and data collection do not systematically cover all of these causes of displacement, let alone the other ways people become internally displaced.

The global figures in this report include only people displaced by conflict and disasters caused by sudden-onset natural hazards. We are working toward global figures for internal displacement caused by development and droughts, but these people remain largely unaccounted for.

We removed figures for Turkmenistan, Uzbekistan and Zimbabwe from our dataset on displacement associated with conflict in 2016, because their primary cause of displacement was forced eviction. This is not to say that the evictions occurred without violence or the threat of it, but because the displacement occurred outside an internationally recognised armed conflict or generalised violence, we stopped counting those affected as IDPs. The scale of this blind spot is significant. In Zimbabwe alone around 266,000 people are currently internally displaced as a result of forced evictions.²⁷¹ They live in near-emergency conditions, at risk of food insecurity and without basic sanitation.²⁷²

We have detailed five ways that people became displaced as a result of the 2013 to 2016 Ebola epidemic in West Africa and the measures put in place to contain the spread of the disease.²⁷³ Given the difficulty of monitoring these phenomena, however, these IDPs are not included in our global baseline.

Incomplete data on the start, dynamics and duration of displacement

From a policy perspective, there are at least two reasons it is essential to have information about specific incidents of new displacement, particularly when it comes to the goal of reducing internal displacement by 50 per cent by 2030. We

need to be able to identify the start of displacement if we are to gauge its duration, which is vital for measuring and addressing protracted situations. If the causes of displacement are to be addressed, it is also crucial to know when, where, how and why new, repeated and secondary displacements occur.

Sometimes we receive information about the number of displacements that have occurred during the course of a particular year. This was the case for Algeria, Burkina Faso, El Salvador, Iraq, Syria and Turkey in 2016. In other cases, such as Mozambique, we inferred the new displacements from the simple fact that there were 15,000 IDPs accounted for in December 2016 against none in 2015.

Without flow data about specific individuals or cases, we use the term “new displacements” to cover the following:

- | People being displaced for the first time from their home or place of habitual residence
- | People previously displaced who had returned or settled elsewhere being displaced again
- | IDPs being displaced from their place of temporary shelter or residence

Given the way data is collected and reported, “new displacements” often mask secondary, tertiary and repeated displacement. In the absence of specific data on each inflow and outflow, we are forced to infer these processes from the contextual analysis of changes in stock data using a consistently applied set of decision rules. In order to avoid miscounting IDPs’ voluntary movements as incidents of displacement we

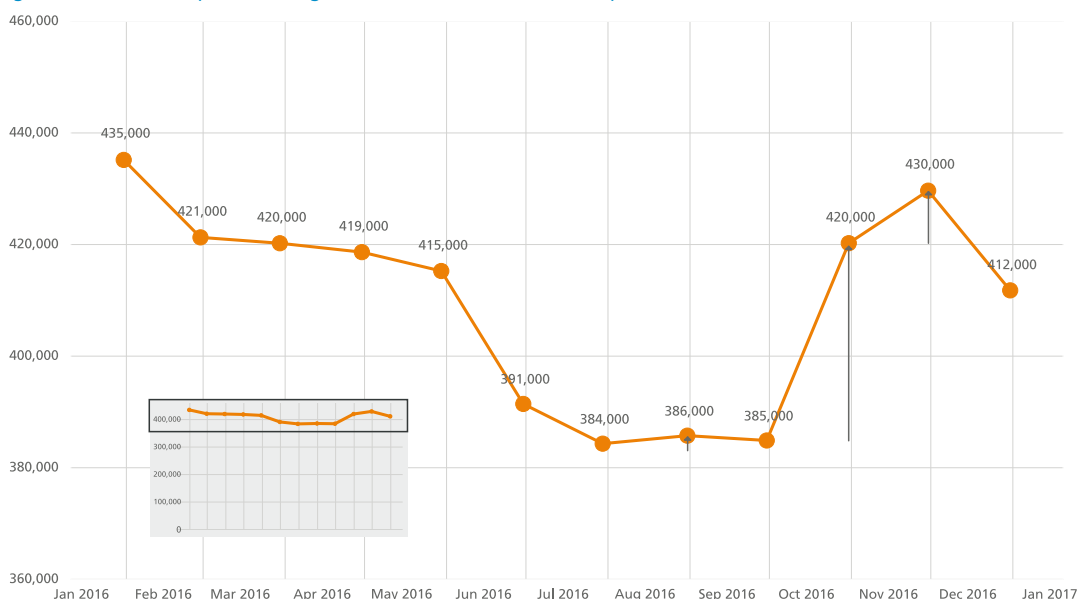
take a conservative line, inferring new displacements associated with conflict when the total number of IDPs in a country increases from one point in time to another, and when the increase is not the result of a change in measurement or methodology (see figure 3.2).

This is an imperfect approach which, depending on the specifics of a given situation, involves varying degrees of uncertainty. For example, based on additional contextual information we have obtained we have reason to believe our estimate of new displacements in CAR to be an underestimate. In the absence of credible quantitative data on the number of new displacements, however, we opt to err on the side of caution and indicate that there were “at least” 46,000 new displacements in 2016 instead of the 60,000 to 80,000 displacements that have been reported to us anecdotally.

Difficulties detecting repeated, secondary and onward movements

We also confront numerous gaps when it comes to covering all flows in and out of displacement. In only a few situations do we receive direct observational data and information about new, secondary or repeated displacement, returns and other processes. Sometimes we receive information about children born to IDPs, which is not the same as new displacement, and deaths in displacement.

Figure 3.2: Internal displacement figures from the Central African Republic



Source: Commission Mouvements de Population

One of the main limitations of working with flow data is that after an initial report of new displacement, there is seldom time-series data that follows up on the evolution of the IDPs' situation over time. Discussed in greater detail in the section below, this lack of follow-up is one of the greatest sources of uncertainty over our year-end figure of 40.3 million people displaced by conflict and violence.

Accounting for secondary and repeated displacement is a particular challenge, as these flows are often recorded simply as "arrivals". Without knowing whether they constitute new, secondary or repeated displacements, we report them as new. The failure to properly capture the dynamics of certain situations makes it hard for those responding, because people who have been displaced for the third or fourth time are likely to face different and greater needs and risks than those who have only been displaced once.

Limited disaggregated data on IDPs' profile and location

This year, we obtained data on displacement associated with conflict disaggregated by sex for 12 out of 56 countries or territories (21 per cent), and disaggregated by age for 11 countries or territories (20 per cent). In other countries some of the data we received was disaggregated in this way, but these datasets either did not cover the entire country or the whole year.

Data disaggregated by IDPs' location and shelter type is also available in only a fraction of the countries we monitor. Displacement data classified as either urban or rural was available in two countries (3.6 per cent), and information reflecting IDPs' shelter type was available in 15 countries (26 per cent).

Without information about who is displaced, where they are located and what kind of shelter they have found, our picture of internal displacement remains one-dimensional. Simply knowing the number of IDPs is not enough for effective programming and policymaking. Those responsible for providing services and protection need to know who their beneficiaries are to ensure that assistance is well targeted and addresses their needs.

ACCOUNTING FOR THE END OF DISPLACEMENT

Quantifying progress toward durable solutions remains a challenge for several reasons. To start with, there is very little data on the different settlement options being pursued, and the IASC framework for durable solutions is an analytical but not an accounting tool.²⁷⁴ There is currently no agreed method, for example, for measuring progress toward durable local integrations in Bogotá, Goma or Kiev consistently. The information required to make consistent assessments is not available either, because data on the necessary indicators is no longer collected or it was never collected in the first place.

A number of experts and institutions, including the Joint IDP Profiling Service (JIPS), the UN Refugee Agency (UNHCR), the International Organization for Migration (IOM) and the UN Development Programme (UNDP) are currently working on ways to put the IASC framework into practice, in order to establish a globally consistent way of accounting for the end of displacement.²⁷⁵ Without such recognised standards and working with the scant information available about IDPs' progress toward achieving durable solutions, we apply a set of decision rules, documented in the methodological annex, to generate consistent estimates across all countries and contexts.

The approach is far from ideal, however, because it may overstate the number of people who have returned, integrated locally or settled elsewhere. Our figures for the three settlement options should not be considered confirmation that the people in question have achieved a durable solution, but rather a statement of what our sources have reported.

Understanding the limitations

SOURCES OF UNCERTAINTY

The process of obtaining data on internal displacement remains a major challenge despite various UN General Assembly resolutions encouraging governments to collect and share their data with IDMC.²⁷⁶ We contact member states every year to remind them that they have requested this of themselves and to offer methodological guidance.

In 2016, as in previous years, some governments – such as Azerbaijan, El Salvador, Georgia and Russia – responded directly by providing some of the data requested (see table 3.1). Others did not respond directly but do collect and publish some data on a regular basis. These include China and the Philippines for disaster and Nigeria and Ukraine for conflict.

Many neither responded directly nor publish data themselves, but work with or allow IOM, OCHA, UNHCR or consortia to do so. These situations often coincide with humanitarian crises or complex emergencies in which international actors are involved. Government involvement ranges from active collaboration, such as Nigeria, to passive involvement, such as Syria.

Table 3.1 also reveals that for some countries our only sources of credible data are civil society organisations, academia or the media. Not surprisingly, our displacement estimates for some of the countries in this group include those where the data is most out of date and for which we have low confidence. We encourage more

countries to either collect and publish this data or, when capacities and resources are a limiting factor, to support the work of others to do so.

One of the most basic but important challenges we face is establishing how to interpret the data from our many sources and map it onto our data model (see figure 3.3). As noted above, the vast majority of internal displacement data is on stocks. Depending on the source, the location may be a specific site such as a camp, a group of sites, an administrative area such as a neighbourhood, city, province or governorate, or an entire country.

Figure 3.3: IDMC's data model highlighting key stocks and flows

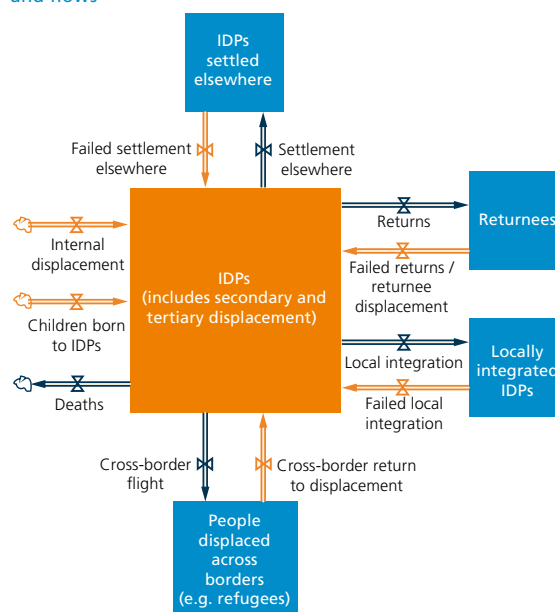


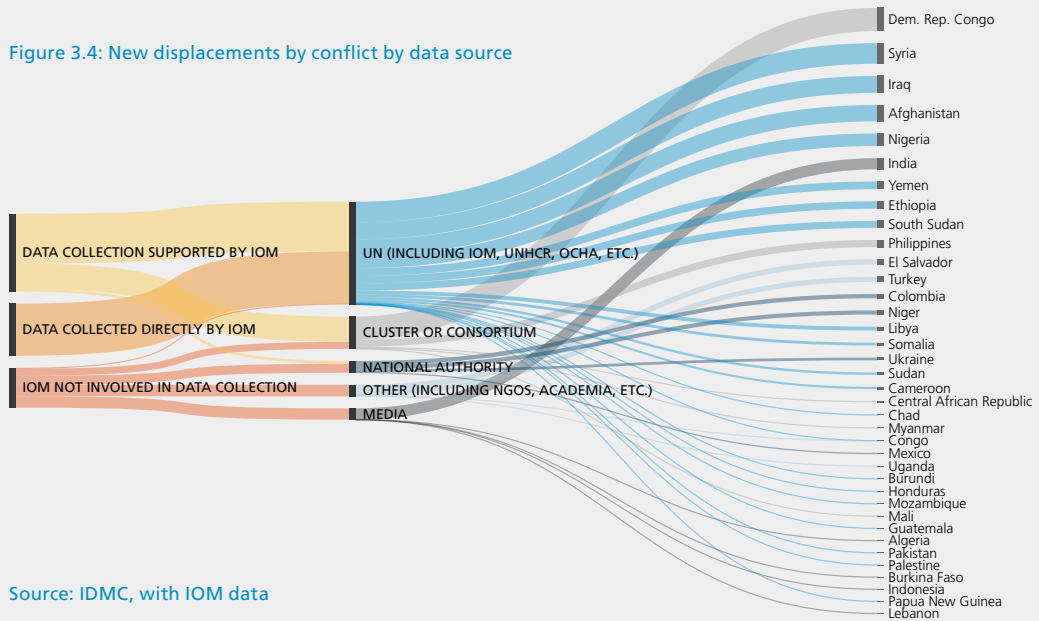
Table 3.1 Internal displacement data sharing scorecard

Government provided data directly to IDMC	Azerbaijan, Congo, El Salvador, Russia, Sri Lanka
Government published data but did not send it directly to IDMC	Bosnia and Herzegovina, Colombia, Georgia, Macedonia, Niger, Peru, Ukraine
Government made data available through a partner (e.g. consortium or UN agencies, NGOs)	Afghanistan, Burundi, Cameroon, CAR, Chad, DRC, Ethiopia, Honduras, Iraq, Kosovo, Lebanon, Libya, Mali, Nigeria, Pakistan, Palestine, Papua New Guinea, Philippines, Somalia, South Sudan, Sudan, Syria, Yemen
No evidence of systematic collection or sharing of displacement data by the government	Abyei Area, Algeria, Armenia, Bangladesh, Burkina Faso, Côte d'Ivoire, Cyprus, Egypt, Guatemala, India, Indonesia, Kenya, Mexico, Mozambique, Myanmar, Nepal, Senegal, Thailand, Togo, Turkey, Uganda

IOM'S DISPLACEMENT TRACKING MATRIX

IOM's Displacement Tracking Matrix (DTM) has become an increasingly important source of data for us. In some countries, such as Iraq, we have collaborated with IOM in designing the DTM questionnaire and methodology. Even when this does not happen, it is the tool used to collect much of the data we analyse, even when we obtain it from another source (see figure 3.4). For example, in the DRC we obtain conflict-related displacement data from the Commission Mouvements de Population (CMP), who in turn relies on the DTM for part of its data collection. The same is true for Burundi, CAR, Mali, Somalia, Syria, Yemen and several other countries.

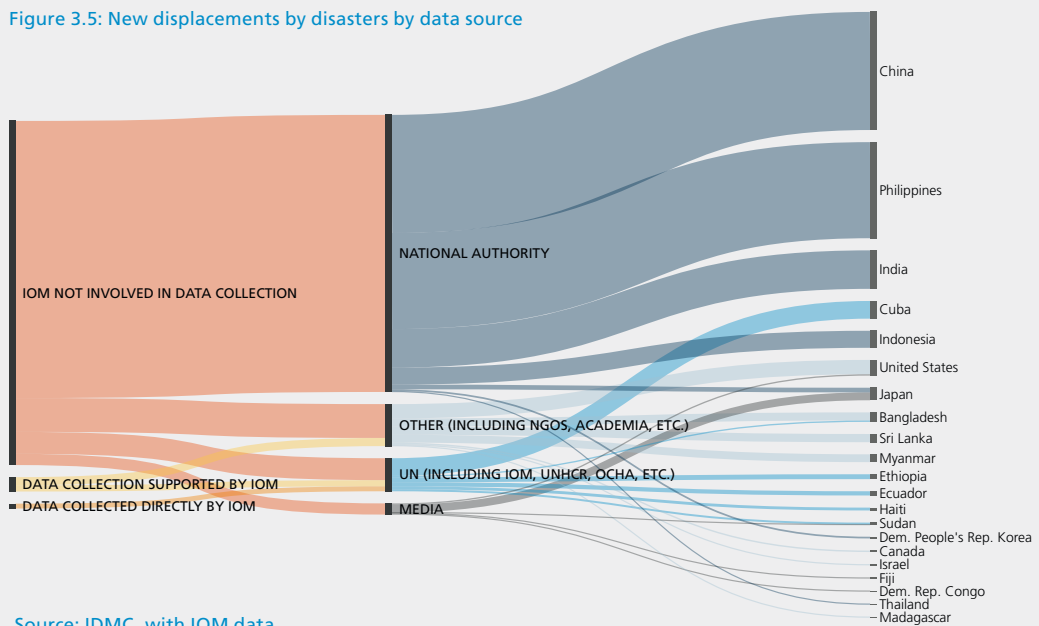
Figure 3.4: New displacements by conflict by data source



Source: IDMC, with IOM data

In the context of disasters, we rely less on IOM's DTM and work more closely with national authorities who collect data and report on events (see figure 3.5). Some notable exceptions include Typhoon Haiyan in the Philippines in 2013 and the 2010 earthquake in Haiti, for each of which IOM collected time-series data over a period of several years. The DTM has played a contributing role even when we receive data from government authorities. For example, the data-collection system used by the Philippines' Department of Social Welfare and Development was based on the DTM and set up with support from IOM.

Figure 3.5: New displacements by disasters by data source



Source: IDMC, with IOM data

Another regular challenge we face concerns the way our sources characterise “returnees”. Sometimes the term refers to IDPs who are said to have returned to their place of origin or habitual residence, but it is also used to refer to people who have returned to their country of origin having fled across a border. In both cases, the people concerned may or not have returned voluntarily, and they are still defined – if not necessarily counted – as displaced until they are able to achieve a durable solution. They may also be displaced again if they find themselves facing renewed violence or another hazard (see part 2).

To make things even more confusing, sometimes data on returnees is needed to measure the number of displacements, rather than the number of people displaced at a particular moment in time. In El Salvador and Iraq, for example, our sources surveyed several hundred thousand people who said they had been displaced by conflict and violence during 2016, but many of whom were no longer displaced at the time they were interviewed. While these people are not included in our year-end figures, we nevertheless need to account for these incidents of displacement.

Why more investment is needed in monitoring over time

One of the key gaps in our data and knowledge concerns what happens to people once they have become displaced. This information is absolutely vital for measuring the extent to which the global target of reducing internal displacement by 50 per cent by 2030 is being achieved. Most importantly, without longitudinal data it is impossible to know how many IDPs there are a given moment in time. This explains why we have struggled to produce stock figures for displacement associated with disasters.

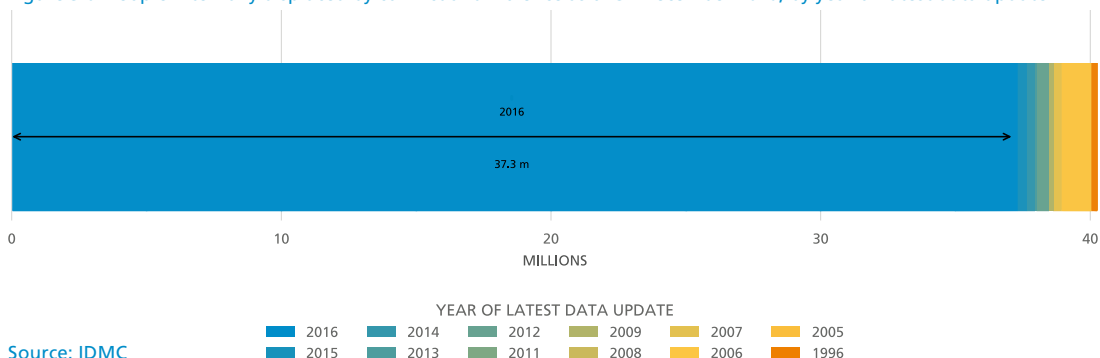
More importantly, halving the number of IDPs is not simply about bringing the numbers down. It’s about ensuring that the people counted achieve durable solutions. Without multiple comparable data sets on different situations and how they have evolved over time, there is relatively little empirical evidence that indicates what works and what doesn’t. Having this information would remove some of the guesswork involved in humanitarian and development financing. It is needed to monitor and evaluate programmes intended to benefit IDPs, to hold those responsible to account and to ensure that limited funding is allocated effectively.

Every time we compile our annual figures for displacement associated with conflict, we find ourselves asking the same questions. Should we use data we believe to be out of date? And if so, how? As we attempted to produce a global stock figure for people displaced by disasters for the first time this year, we confronted the same problem in a different form. What do we do when data stops being collected before the number “goes back to zero”? Do we keep counting these IDPs? And if so, for how long before we feel that the data no longer accurately reflects the situation on the ground?

Outdated information on certain conflict caseloads

Last year, we introduced the concept of “decaying” data and presented our estimates for displacement associated with conflict based on the age of the source data. We have done so again this year. Around 93 per cent of the people displaced by conflict as of 31 December 2016 are accounted for by data that was last updated during the year (see figure 3.6). That said, a small number of our estimates are based on old data sources, some of them more than a decade old.

Figure 3.6: People internally displaced by conflict and violence as of 31 December 2016, by year of latest data update



Source: IDMC

Much of the data we rely on is collected by UN agencies such as OCHA, UNHCR and IOM and their NGO partners. In some cases, it is collected by an institution or consortium mandated with that single primary function, such as DRC's Commission on Population Movements, the Task Force on Population Movement in Yemen and IOM's numerous DTM operations.

Much of the time, however, data on IDPs is collected by institutions working under broader mandates, such as the UN's humanitarian profiles, humanitarian needs overviews and humanitarian response plans. In these cases, the data is updated only a few times a year, and often lapses once the humanitarian phase of a crisis has ended, even if the displacement has not.

As noted in part 1, our source of data on internal displacement associated with conflict in Colombia comes from the government's registry of victims (*Registro Único de Víctimas*, RUV). The purpose of the RUV is to account for all victims of the conflict. This involves identifying people who are or were internally displaced, but it does not necessitate tracking them over time. Once someone has been registered they remain so, meaning there is little or no follow-up information with which to determine whether or not they are still displaced.

Gathering time-series data systematically can be costly and sometimes a lack of funding means collection falls off before a crisis is resolved. When various crises compete for attention and resources, some inevitably lose out and become neglected, which translates into less funding and political will to stay on top of them. This can occur even when the number of IDPs is significant, as has been the case with Burundi, where more than 141,000 people were displaced at the end of 2016.²⁷⁷ For most of the year, IOM's DTM covered only three of 18 provinces, excluding Bujumbura Mairie, one of the locations most affected by internal displacement in the country. In September, IOM's DTM coverage expanded to seven provinces and in December to 11 (although still excluding Bujumbura).

At the end of the year OCHA published its annual Humanitarian Needs Overview for Burundi which also included IDP estimates for Burundi. OCHA's figures differed from IOM's in that they covered all provinces in the country, were collected at different intervals and were based on a different estimation methodology that placed more emphasis on expert opinion. In the few cases

where both institutions published displacement figures for the same region, these differences resulted in widely disparate estimates, generating even more confusion and casting some doubts on the accuracy and reliability of both datasets. For example, for the month of July 2016 IOM estimated that there were 2,444 IDPs in Rumonge province compared to 13,095 reported by OCHA.

Longitudinal data collection also ends, or is interrupted, based on government policies. This can involve who is counted as an IDP or where data collection is undertaken or permitted. These are common challenges when a government has adopted a policy that specifically aims to reduce the number of IDPs, as in Kenya, or when it wants to shift attention away from a particular crisis.

As in last year's report, we have included these "decaying" figures, which were last updated prior to 2015 (see figure 3.6). This year, these figures account for only 6.5 per cent of the global total. We publish information about the age of the data for two reasons. It allows readers to draw their own conclusions about the figures, and by depicting them in this way we hope to encourage anyone with more recent data to come forward, or to help follow up on the situations in question if no more recent data is available.

Next year, we plan to remove the following figures from our global total unless we receive updated information (see table 3.2).

For a more comprehensive and transparent assessment of our confidence in the data we have provided, please see the methodological annex at the end of this report, where the age of the data, its geographical reach and other factors are further discussed and evaluated.

Table 3.2: Caseloads to be removed in 2018 unless we receive updated information ^{278 279 280 281 282}

Country	Figure	Date of latest data update
Guatemala	242,000: this is based on an estimate published by the United Nations Population Fund (UNFPA) in 1997 for people displaced by the civil war (1960-1996). A 2000 report produced by FUNCEDE, Fundación Arias para la Paz y el Progreso Humano and UN-HABITAT (“El Fenómeno del Desarraigo en Guatemala”) disaggregates the 1997 UNFPA Figures on which IDMC’s 2016 estimate is based: 242,000 IDPs left by the armed conflict in 1997. ¹	31/12/1996
Armenia	8,400: this is a decaying figure which has not been updated since 2005. Of 8,400 IDPs, 2,600 are from the enclave Artsvashen and have no realistic opportunity of returning to their former place of residence. A further 2,480 are potential IDPs, whose current status and whereabouts are unknown. The remainder is staying in different parts of Armenia and has not managed to find durable solutions. This figure comes from a profiling exercise led by NRC and the armenian State Migration Service in 2004 (IDP Mapping Survey 2002 - 2004). ²	31/12/2005
Bangladesh	151,000: this is a camp population figure for the Bihari population, issued by the local NGO Al-Falah. Al-Falah was hired by UNHCR in 2006 to conduct a profiling study of the Urdu-speaking/Biharis living in informal settlements since their postwar displacement in the 1970s. Correspondence between IDMC and Al-Falah in 2016 and 2017 has confirmed that there has not been a similar exercise to update the IDP subpopulation profile. ³	30/05/2006
Bangladesh	275,000: this figure corresponds to internal displacement in the Chittagong Hill Tracts, and stem from a 2009 socio-economic baseline survey of the Chittagong Hill Tracts, by the Human Development Research Centre; the displacement figure at the time of the survey came to 275,000 people. Some evidence was provided in March 2017 that this figure has come down, although it could not be verified to a satisfactory standard. ⁴	31/12/2007
Turkey	954,000: this is an uncertain figure based on the most recently available data on displacement in Turkey. It is based on a 2006 study commissioned by the government and carried out by Hacettepe University. The study concluded that between 954,000 and 1.2 million people were internally displaced from 1986 to 2005. The study found that IDPs had fled their homes due to armed conflict between Turkey and the Kurdistan Workers’ Party (PKK). More than 30 percent of the persons displaced by violence fled their homes between 1986 and 1990, 60 percent between 1991 and 1995, and the remainder between 1996 and 2005. ⁵	06/12/2006

Disasters: difficulties understanding displacement patterns over time

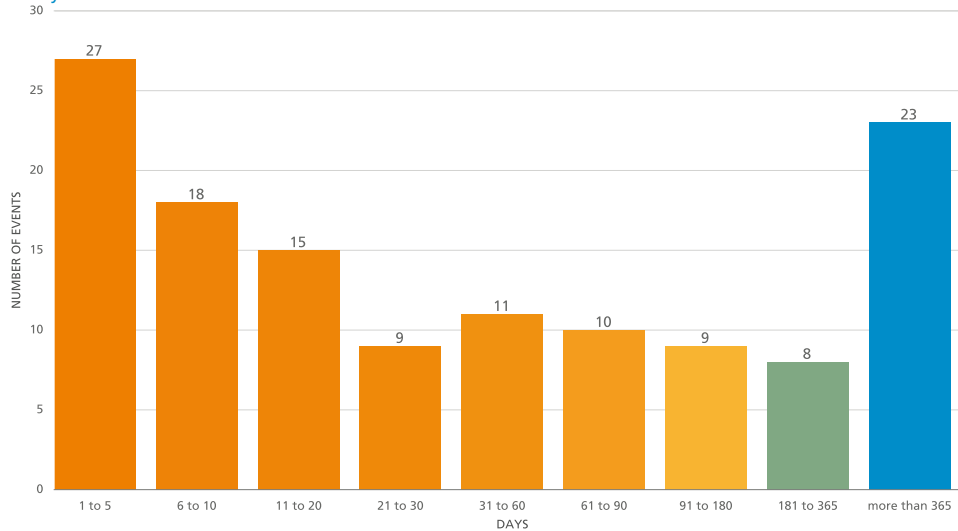
We provide an annual global stock figure for people displaced by conflict and violence based on the best data available as of the end of the year, but as yet we have been unable to do so for people displaced by disasters.

In 2016, we began working to address this.²⁸³ The lack of information available on displacement associated with specific events over time, time-series data, is a serious limitation to this exercise. In order to present our first estimate

of the number of people living in displacement following disasters as of 31 December 2016, we collected as much time-series data as possible for disasters that caused the 50 largest displacements in 2016 and the ten largest each year from 2008 to 2015.

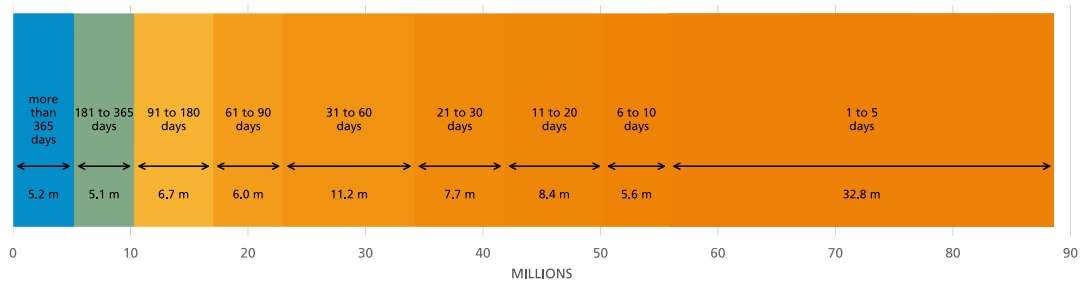
This sample consists mainly of hydro-meteorological disasters such as floods and storms. These disasters make up 86 per cent of the sample and tend to predominate among the largest new displacements each year, but they may not cause long-lasting displacement compared to earthquakes. The period of time for which data was collected following each disaster varied considerably (see figure 3.7).

Figure 3.7: Length of time for which data was collected following the 50 largest displacements in 2016 and the ten largest each year 2008 to 2015



Source: IDMC

Figure 3.8: The magnitude of the estimated stock figure for displacement by disasters, based on the length of time for which data was collected



Source: IDMC

For more than half of the sample, the “last” data point came within a month of the disaster striking, and for several it was more a matter of days. Forty events, less than a third of the total, yielded data recorded three or more months later. Only for around 18 per cent was displacement still reported on after a year. Given that many of these displacement events would have been accompanied by the widespread damage or destruction of homes, livelihoods and basic infrastructure and services, we would expect more of them to have involved displacement continuing for over three months.

Underscoring this point, when we examine time-series data for individual disasters we see that data collection almost always ends before the number of people displaced returns to zero (see Philippines box below). Nor does available data account for multiple displacements of the same people within the same year, if they return only to face ongoing risks, or are confronted with similar or new risks that cause their secondary or onward displacement from their place of refuge.

The lack of monitoring over time and the limited availability of longitudinal data has significant consequences in estimating the total number of people displaced at a given point in time. Compared with data collected several months after a disaster, that collected immediately after it strikes tends to relate more closely to the peak number of people displaced, particularly when this data relates to the number of people evacuated. This obscures the fact that a significant number of people may have been displaced for days or weeks rather than months or years.

This is the case when voluntary returns occur and when the homes of evacuated people are not destroyed, rendered uninhabitable or remain inaccessible. If we were to include all events from our sample for which only peak displacement data is available, it would result in a figure of around 89 million people (see figure 3.8). That is an overestimate because it does not account for everyone who returned home or established a new home elsewhere after data collection stopped.

If we apply more stringent criteria and only include events for which there was data for at least three months following the onset of the disaster, we would arrive at a figure of around 17 million. This of course excludes any displacements that occurred in the last three months of 2016, but given that it is also based on data that is out of date in some cases, it may also be an overestimate – albeit a less significant one.

At the same time, the partial nature of the data available for most events means that some displaced people, particularly those in protracted or chronic situations, are currently off our radar screen. The data is simply too scant to allow us to gauge the overall situation with any accuracy. Having said that, updated data indicates that at least 3.6 million people remain displaced due to three large disasters in recent years – the earthquakes in Haiti (2010) and Nepal (2015), and Typhoon Haiyan in the Philippines (2013).

Given the considerable difference between an estimate of 89 million and 17 million, and the fact that our analysis is based on only 130 out of more than 2,000 events in our dataset, we have decided not to endorse either figure or anything in between. Instead, we have described the diffi-

culty we faced as a clear call for more follow-up and data collection on people displaced by disasters.

As our previous reporting has shown and as our displacement risk modelling suggests, many people are exposed and vulnerable to frequently occurring, low-intensity hazards (see part 1). This increases the risk of repeated displacement and undermines long-term development gains for these communities.

Despite our lack of confidence in stock figures for displacement associated with disasters, our analysis of the time-series data available serves several purposes. It provides a useful stocktake of how much data has been captured, for how long and by whom. This has led to the identification of some good practices (see Philippines box below).

It also sheds light on patterns of displacement for different hazards, which helps explain how we have estimated the total number of new displacements for those events. Flooding in the Indian state of Bihar in July and August 2016, for example, generated two distinct waves of new displacements which resulted in 1,670,000 displacements.

A group of children play near the edge of the IDP camp on a dirt road in Khmer, Amran, Yemen. Photo © UNHCR/Rawan Shaif, February 2016

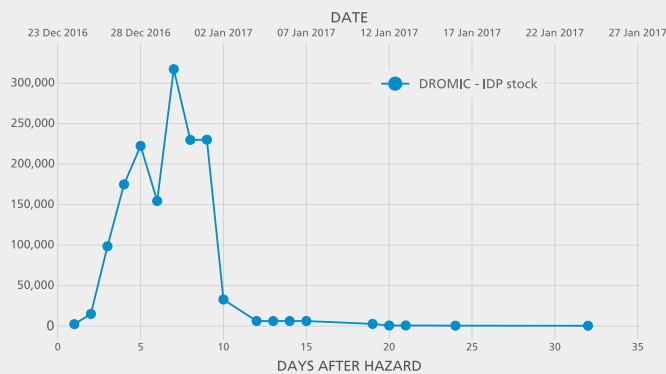


PHILIPPINES: A MODEL FOR CAPTURING AND REPORTING TIME-SERIES DATA

Thanks to its strong law and policy on disaster risk reduction and management, and its frequent and extensive experience in responding to disasters, the Philippines does a better job than many wealthier countries of collecting and sharing data on disasters and the displacement they cause.

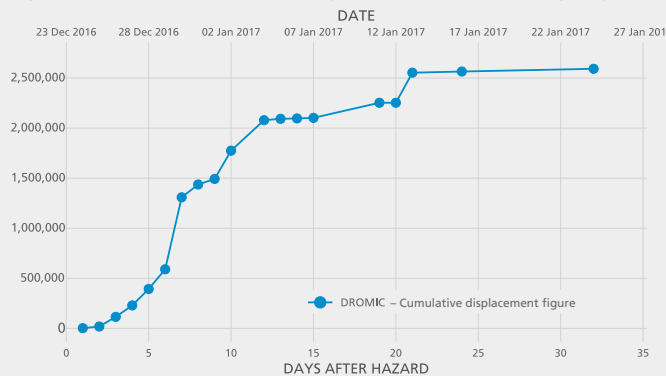
Its National Disaster Risk Reduction and Risk Management Council (NDRRMC) and Department of Social Welfare and Development (DSWD) publish situation reports for several days after each disaster, and twice-daily reports for the first nine or ten days after large ones. The reports include the number of people displaced to official evacuation centres and elsewhere at a given moment in time, and a running tally of the number of people a disaster has displaced over time (see figures 3.9 and 3.10). Our figures for typhoon Nock-Ten, for example, are based on 24 situation reports published over the first few weeks of the disaster.

Figure 3.9: Displacement data for typhoon Nock-Ten (locally known as Nina), number of IDPs at a given moment in time



Source: DSWD/Disaster Response Operations Monitoring and Information Center (DROMIC)

Figure 3.10: Displacement data for typhoon Nock-Ten, running tally of people displaced over time



Source: DSWD/Disaster Response Operations Monitoring and Information Center (DROMIC)

The daily stock figures reveal a spike in displacement on 30 December, eight days after the typhoon struck, but the cumulative figure continues to rise for several more days. This is because the data collection process identified additional returnees, people who had been displaced but had returned by the time they were counted. The challenge when dealing with these two sets of time-series data lies in bringing them together in a logical and methodologically sound way. The cumulative figure, for example, tells us how many people were displaced, but it doesn't tell us when, where or for how long.

Of all the time-series data we obtained, in only five of the more than 130 displacements did collection continue until the number of displaced people reached zero. Two were in the Philippines, and the others were in Indonesia, Tonga and India. This represents a major blind spot, with significant implications for people who remain displaced but not counted, and those responsible for protecting them. The fact that data collection ended while people were still displaced in more than 130 displacements further underscores the need for much greater investment in monitoring displacement over time in all countries.

Painting a fuller picture

INNOVATIVE SOLUTIONS

Given the importance of accurate information on new displacements and the evolution of situations over time, we have begun to incorporate new approaches to our monitoring (see table 3.3). Taken together, our new “hybrid” approach combines event detection and data collection with the analysis of time-series data when it is available.

For displacement associated with conflict, we have begun to identify and capture data about incidents of new displacement manually. In order to address the challenge of event detection on a global scale, we are also developing a new semi-automated process to identify potential displacements for human verification (see p.84).

For disasters, we already capture several hundred incidents of new displacement a year – good but still not global. We tend to miss displacements associated with localised disasters that affect small numbers of people. The bigger gap, however, is in the systematic collection of time-series data on people once they have become displaced. We have begun working with partners to collect and analyse more of this data so we can infer both the total number of people displaced by an event, and track the number and needs of displaced people as they evolve over time.

One such method involves analysing satellite imagery to detect changes in human habitation in response to development projects such as dams, natural hazards and conflict. Based on the number of buildings destroyed or the extent of flooded land and population and settlement data, we will infer how many people may have been displaced, an approach already used by our sources to triangulate data obtained from the field.

Another approach will transform our probabilistic risk model for displacement associated with disasters (described in part 1) into a real-time tool to support monitoring. When a hazard has been detected or is predicted to occur, we will simulate the amount of destruction and displacement expected to result.

Using satellite imagery analysis and our displacement risk model as monitoring tools will help us extend the geographical coverage of our monitoring and address some of the factors responsible for the incomplete picture of displacement, notably language, reporting and selection biases.

Table 3.3: Challenges and solutions for our “hybrid” monitoring approach

Context	Current situation	Way forward
Conflict	Time-series data with a focus on end-of-year updates of stocks, but limited event detection for new displacement	Systematic event detection to inform the collection of time-series data and more data points over time
Disasters	Event detection with a focus on the occurrence of new displacement, but limited time-series data	More comprehensive event detection and systematic collection of data about how displacements evolve over time

#IDETECT

In January 2017, IDMC and the UN launched the Internal Displacement Event Tagging, Extraction and Clustering Tool (#IDETECT) challenge on the UN's data science crowdsourcing platform, Unite Ideas. It has brought together teams representing dozens of data scientists from around the world to develop a new tool that we will use to monitor displacement associated with disasters, conflict, violence and development projects.

#IDETECT will expand and diversify the sources we use for monitoring significantly, helping to address – though not eliminate – some of the factors that impede our painting a comprehensive global picture. The tool will cast a wide net so we can obtain information about, analyse and shed light on far more displacement situations than we currently do (see figure 3.1). That said, #IDETECT's scope will still be limited to events reported in the media or by partners in the field. To overcome this reporting bias, we have also begun exploring further approaches to detect displacement using other types of data and means of analysis.

The tool will make our monitoring more efficient and comprehensive, and it will also provide the humanitarian community with an easy way to extract and analyse facts from any type of documents, be they news, field reports, social media or other sources.

HOW IT WORKS: FILTERING AND TAGGING

The first step is to mine huge datasets of news, such as [the GDELT Project](#), the [European Media Monitor](#) and social media platforms, and extract records that relate to displacement. The next is to tag the events as being related to conflict, violence, disasters or other cause or trigger.

NATURAL LANGUAGE PROCESSING

The tool will use natural language processing (NLP) to extract certain facts from the source material including, but not limited to:

- | The publication date of the document
- | The place where the displacement reportedly occurred
- | The number of people displaced

DATA VISUALISATION, HUMAN VALIDATION AND MACHINE LEARNING

It will then visualise the data for us and our partners to review. The results of this human validation process will inform the NLP so that it performs more accurately in the future, a process known as supervised machine learning (see figure 3.11).

Figure 3.11: Visual representation of the #IDETECT tool and validation process

