

Kosovo Memory Book Database Presentation and Evaluation

February 4th and 6th 2015
Pristina, Belgrade

Speakers:

Nataša Kandić, Coordinator of the Kosovo Memory Book Project
Patrick Ball PhD, Executive Director of the Human Rights Data Analysis Group, dealing with truth and reconciliation commissions, through database development and data analysis
Professor Michael Spagat, Head of the Economics Department at Royal Holloway, University of London

Nataša Kandić:

I am going to talk about how we created the Kosovo Memory Book database - and by “we”, I mean the Humanitarian Law Center (HLC) and the Humanitarian Law Center Kosovo (HLCK) teams. Dr. Patrick Ball and Prof. Michael Spagat will then present their findings and conclusions about the reliability, comprehensiveness and completeness of this database. Both of them have been conducting qualitative and statistical analyses of human rights violations in armed conflicts for many years, and so our database was placed in the right hands. Two days ago, all of us here gave a presentation on the KMB database in Prishtina, in the amphitheater of the National Library. The event was attended by more than 200 people, and everyone was pleased to hear me say that this database of victims in Kosovo stands as a truly unique document, since it has been created jointly by both Albanian and Serbian researchers, willing to cross all national divides and borders in the aftermath of the wars and war crimes, and highly motivated to record all the facts and victims, and thus contribute to the creation of a shared history of the events of the nineties. The Humanitarian Law Center began its work on documenting human rights violations and war crimes as soon as it was established in November of 1992. Soon after, we opened an office in Kosovo. The HLC’s researchers were already in the field when the armed conflict in Kosovo broke out in February of 1998. In 2010, our Kosovo office grew into an independent NGO.

As I said, the database of human losses in connection with the war in Kosovo between January 1st 1998 and December 31st 2000 was created jointly by the Humanitarian Law Center and the Humanitarian Law Center Kosovo. It contains witness and victim statements, statements by victims’ families, official court documents, books, media reports and numerous other documents. It currently holds a total of 31,600 documents, 14,612 of which are statements taken from surviving witnesses and victims’ relatives. All of the records contained in the database are verifiable and based on empirical research. In the course of our field work, we have gathered over 12,000 statements concerning 14,000 individual victims, including not only the true victims, but also those with respect to whom there are some indications of victimhood. We verify all cases of possible victims in the field, so that not a single true victim is left out. I do not believe that it is possible to create a database that would include every single true victim, but our goal is to investigate every report and do everything in our power, at least in the investigative sense, to ensure that not a single victim is forgotten. Our research is actually based on gathering records using three basic methods: the first method involves gathering statements, which I have already spoken about; the second method involves gathering documents and analyzing their content, particularly the content of

documents containing personal victim information and court documents describing the circumstances of their death or disappearance; and the third method involves the verification of all the records gathered. This stage of our research is still underway.

So, what exactly have we done? Once we had gathered about 12,000 witness statements through our research, we decided to compare our victim data with data from other sources, and we selected four major registries of ethnic Albanian victims for this purpose. Two of these sources, the Kosovo State Archives list compiled by Jusuf Osmani and the list compiled by the Kosovo Democratic Alliance in 2002, each contained the names of more than 10,000 war victims. The third major list, containing a little over 9,000 names, was compiled by the Norwegian humanitarian Josef Martinsen, as part of a book published in English. The fourth list reflects records obtained by the Kosovo Council for Human Rights (a Kosovo-based NGO that has been documenting war crimes and other human rights violations for years) and contains the names of more than 8,000 victims. The preparations for the data verification stage began in 2012. In order to be able to compare the records, we first had to input all of the data contained in the four selected registries into our database, which proved to be a very time-consuming process, and required a whole team of data transfer experts. The data comparison revealed that these four registries contained the names of 3,190 victims not present in our database. Since the goal of our research was the creation of a complete registry of all true war victims, we then decided to begin conducting additional research into the status of all reported victims not already registered in our database, while, at the same time, continuing to research and document existing war victims and the circumstances of their death or disappearance. By accounting for spelling errors resulting from the mixing of languages and alphabets, we were able to match about 10,000 victim names from these sources to those already registered in our database. Before the war, most names in personal documents used to be spelled in Serbian, but once the war ended, people started to spell their names in their native languages, even before they were issued with new IDs. Because of all this, we encountered significant differences in the spelling of names which contained the Albanian letters “glj” and “xh”. At the same time, verification in the field revealed that 488 of the 3,190 victims reported in these sources were still alive, and had therefore been reported erroneously. We also established that, among those listed, 151 persons died a natural death, while 184 persons who did die in the war, actually died owing to causes completely unrelated to the war. In total, 581 persons were found to be true war victims. With respect to 715 persons listed as victims in these four sources, no records could be found that they had ever existed, despite multiple inquiries having been made with municipal registrars, local officials and boards of political parties in the places of their alleged residence or birth. We also discovered cases of victims having been registered multiple times under two or three different names. In addition, we established that the names of more than 300 victims were mentioned only once in sources such as books and publications by various authors, associations and private persons, with no source of victim information listed. We discovered that 250 Albanian and Serbian victims were mentioned only in media reports, also with no source of information listed. To find a person just by their name, who has been mentioned only once and in one source only, is a hard and often impossible task. But, it is still part of our goal of not allowing even a single victim to be forgotten.

In the end, the 31,600 documents we have gathered confirm, undeniably and unambiguously, the death or disappearance of a total of 13,535 persons. You can see each of the 13,535 victim names displayed on this monitor here, along with their complete personal records, including whether they were civilians or combatants. We have established that, in 1998 alone, a total of 2,156 persons died or disappeared in connection with the war. With respect to civilian victims only, 1,100 of them were Albanians, 132 were Serbs, while the remaining 46 were of Roma and other non-Albanian ethnicities. With respect to military casualties, we

have determined that a total of 703 members of the KLA and 175 members of the Serbian MUP or VJ died in connection with the war in 1998. From January 1st until June 14th of 1999... but let me first say that we have chosen June 14th as the last day of the conflict, because a crime in which civilians died was perpetrated on that day, and we thought that it would be important to include these victims as well... well, during this period, and I am talking about a period including the duration of the NATO-led air campaign, a total of 10,120 people lost their lives or went missing in connection with the war, 758 of them having died in NATO's airstrikes. The Serbian MUP and VJ were responsible for the deaths of around 7,000 Albanian civilians, while the KLA was responsible for the deaths of 200 civilians of Serbian ethnicity. 172 Roma civilians also died in the same period. Between January and June of 1999, a total of 1,354 members of the Kosovo Liberation Army and 860 members of the Yugoslav Army and the Serbian MUP were killed. The post-conflict violence refers to the period after the withdrawal of SRJ security forces and the arrival of international forces in Kosovo, up until December 31st, 2000. We have selected this period by monitoring the frequency of disappearances and killings of ethnic Serbs, Roma and other non-Albanians, including the so-called 'loyal' Albanians. We have established the first and last names for a total of 1,257 persons who lost their lives or disappeared in connection with the war during this period - 717 Serbs, 307 Albanians and 233 Roma and other non-Albanians. What remains to be done, is to complete the verification of the data and status for about 2,000 victims not included in our registry of true war victims. Other than that, we have yet to publish the remaining three volumes of the Kosovo Memory Book, containing narratives about each individual victim.

What would be some other good reasons to use this database? Let me first refer you to the trial of the members of the *Scorpions*, accused of executing six prisoners from Srebrenica. This database contains a video recording which was considered to be one of the most important pieces of evidence in these proceedings. It has already been used in the course of two such trials, and it was owing to this recording that two members of the *Scorpions* were convicted in Serbia over the killing of six Muslims from Srebrenica, while another member of the unit was convicted in Croatia. This database has become a sort of a challenge to all the researchers working on human losses, and through the Kosovo Memory Book website, we receive daily messages and requests to use the database for scientific purposes. Thanks to this database, the Humanitarian Law Center has created dossiers about the 549th, 125th and 37th Mechanized Brigades of the Yugoslav Army, and the *Diković* and *Rudnica* dossiers. This would be a good opportunity to stress that all NGOs working on documenting war crimes have an obligation to speak out, in public, whenever they encounter any indication that a civilian official or a highly ranked officer in the army or the police was involved in the ordering or perpetration of war crimes, or that war crimes were perpetrated in the territory within their jurisdiction, and a duty to expose them to public shaming and curiosity, and demand that they be removed from office and prosecuted. We now find ourselves in a privileged position, because the Hague Tribunal, unlike the states of the former Yugoslavia, has left us a rich legacy – the documents created by the region's armies and police forces, which no NGO would have been able to obtain on its own by filing requests with the authorities. But now, thanks to the Hague Tribunal, these documents are publically accessible - not just to NGOs, or the Humanitarian Law Center and the Humanitarian Law Center Kosovo, but to local war crimes chambers as well, and to all citizens who wish to judge for themselves what happened in the past.

And, what I think is most important, and I leave this to the last on purpose, is the fact that the Kosovo Memory Book is part of a larger project - a larger event, that is. It does not concern only Serbia, but all the states that took part in the wars, starting from 1991. A huge coalition has now gathered around the idea to establish a Regional Commission which, in turn, would

establish all the facts about the war crimes and victims of the wars fought in Croatia, Bosnia and Herzegovina, Kosovo, Slovenia and Macedonia. This idea of ours has received a good deal of public support, but also tangible, political support. A year ago, the Presidents of all the states in the region except Slovenia delegated their representatives to RECOM, and they have been reviewing the RECOM Statute all year long. The representatives did make some changes to this document, created by a regional coalition composed of 2,000 civil society organizations, and in November of 2011, the Coalition for RECOM gave its support to their amendments, seeing them as a significant improvement over their original Statute proposal. However, the gap between the civil society and institutions of state is sometimes so wide, that it seems impossible to bridge. The civil society organizations are always in the position of having to beg for a meeting, and to rejoice should the secretary say, "He's not in, try again tomorrow." And the following day, the answer would be the same. Sometimes, we have to call for weeks on end just to schedule a meeting with some representative of some state institution. Apart from all that, and bearing in mind that this is a regional initiative, we dread the outcome of every election in the region, and there is at least one held every year - sometimes parliamentary, sometimes presidential, and sometimes both at the same time. In practice, this means that the support of one president lasts only as long as his mandate. With the election of any new president, the Coalition for RECOM finds itself back at the beginning. Professor Žarko Puhovski, one of RECOM's public advocates, says that, with the election of Kolinda Grabar Kitanović, he will have to start building anew what he has already accomplished at least three times by now.

Towards the end of this month, we will hold a press conference in Sarajevo and give a presentation on another project of ours, aimed at documenting prison camps and other places of detention in Bosnia and Herzegovina. If RECOM ever gets established, that too would be one of its tasks. With projects like these, NGOs are not only expanding their frontiers of cooperation with state institutions, but also providing a direct contribution to the delivery of justice for past crimes, as well as the creation of a shared social memory. I would like to firmly believe that the support of the Serbian President and Prime Minister for the establishment of a Regional Commission is strong, and not subject to changes influenced by daily politics. We need to know the name of every victim of war, irrespective of the number of victims on the other side. In death, we are all equal, and we need this for our own sake, and the sake of others. We also need this for the sake of future generations, so that we can finally put an end to all the manipulation, because numbers are the easiest to manipulate. I want to believe that someday we will be in a position to say that, despite all obstacles, criticisms and personal attacks, we were able to identify the people who lost their lives in these wars. I am now going to give the floor to Dr. Patrick Ball, who will present his findings and conclusions about the Kosovo Memory Book database, and give us some recommendations.

Dr Patrick Ball:

This is a very important, and very valuable, very high quality database. And, let me tell you why I think that, and on what grounds I make that conclusion. I have been looking at data on human losses for 24 years, and I have done this kind of work, building databases and conducting statistical analyses of the pattern of mortality among the victims in over 30 countries. I've worked for four international criminal tribunals, two national war crimes tribunals, four UN missions, nine truth commissions, and many, many, many non-governmental organizations all around the world, and I've looked at hundreds of databases

during that time. In my professional opinion, this is one of the very best of the databases that I have seen of this kind. It is reliable and complete, or quite nearly complete. And, so I will give you some small insights into why I think this. Our complete report is very, very detailed and technical. It is available on the Kosovo Memory Book's website, so you can download the complete report and consider it in all of its details. But, I want to give you today just very high level, very specific examples about why we have reached this conclusion that this database is reliable and complete. First, we visited the Humanitarian Law Center, my colleague Julle Krüger and I, visited four times to the Center, and we interviewed the people in the team who were conducting the research and managing the database. We examined the technology of the database, we did descriptive analysis, statistical analysis of the database, we analyzed the documentation per victim, and perhaps most importantly, we compared the HLC's to 10 other databases of human losses in Kosovo during the relevant period. I want to give just very specific examples about how we reached our conclusions. First, we think it's really important to understand that each of the reported deaths has many sources of information, and I'll talk about that in a bit more, a bit more detail. Natasha has already explained how each of the deaths in the database has already been investigated individually in great detail. These are not rumors, but rather confirmed investigations that have been done quite professionally. And, based on our analysis, comparing the database of the Kosovo Memory Book to other databases, we believe that nearly all of the victims of the period have been recorded in this database. So, let's look first at the analysis of documentation for each victim. Each of these three graphs shows the number of victims who have a certain number of sources. So, in the graph in the upper left, which has a kind of up and down shape, you see that the highest point of that graph is at approximately 10 sources to the horizontal axis's number of sources per victims. That suggests that the most frequent number of sources per victim is 10 – 10 independent sources of information for each victim on average in this database. That is an extraordinary amount of detail on each of the individual cases. I have never seen a database so intensely well documented as that. Notice that there are some victims that have 15, 20, even 25 independent sources of information, for each victim. I find that quite extraordinary. It is useful and interesting to compare the graph in the upper left with the other two graphs, which are also presented here. Graph B in the upper right shows the number of sources for the potential victims, that is, victims whose names are known to the Humanitarian Law Center, but are not confirmed yet as victims of the war. Notice that there are many fewer sources for the victims that are not confirmed. This may seem quite reasonable, because it takes more sources perhaps to confirm, but we believe that the comparison between these two graphs is one way of understanding, at a purely statistical level, the very intense level of detail and rigor that the Humanitarian Law Center has brought to this project. Finally, in the lower graph, on the bottom row, we see the number of sources for the victims who have been confirmed as “not war victims”, that is, victims of accidents, or of disease, or of some other cause that is not related to the war. There is more than one source, on average, for each of these victims, but not nearly as many as for the other victims, because it takes less information to confirm that this person died by accident, or by disease. So, this is one of our several kinds of analyses in our report which lead us to conclude that the information in the Kosovo Memory Book database is reliable. But, there is a second question that, to statisticians, is also very interesting, which is – is the database complete, that is, are

there victims who are excluded from this database? And, we believe there probably are a few victims, but not very many. This is a very complicated table, and I'm only going to interpret the tiny bit in the very far right corner of this table. If you are interested in all the detail, I'll refer you to our report. This is table 6.2. But, the point of this table, the point that I want to bring to your attention in this brief discussion this morning, is that the five sources in the bottom five rows of this report were not available to the Humanitarian Law Center when they did their work. These were sources that were available to me as, in my capacity as an expert for the Hague Tribunal, as well my previous work before I became an expert before the Tribunal. These are other databases of human losses in the conflict, but databases, and this is quite important, that the Humanitarian Law Center did not have access to. And, this is useful because, having these databases, that the HLC had not seen before, we were able to compare to the HLC's database to determine if these databases contained victims that the HLC did not document. And, as you see in the final column, in the far right, there are only a very tiny number of victims who we found in these other databases, that the HLC had not already documented in all the detail that I showed in my previous slide. This is quite extraordinary. It's a tiny number of victims that were missing from the HLC's list. We therefore conclude that there is only a tiny number, a very small number of additional victims who have not yet been documented by the HLC. As Natasha mentioned, surely there are some, but the number would be quite small. To conclude, this database offers a reliable and nearly complete enumeration of the victims of the conflict in Kosovo from 1998 to 2000. I recommend it strongly, and without reservation, as a foundation for future academic analyses, scientific analyses, for historical memory, to remember and validate the memories of those who have been killed, as well as for transitional justice mechanisms and any further accountability. I congratulate the HLC on this extraordinary achievement, and the HLC Kosovo. Thank you, Natasha. I congratulate both groups on this extraordinary accomplishment, and it has been my honor to evaluate it during our work together in the last two years. Thank you very much.

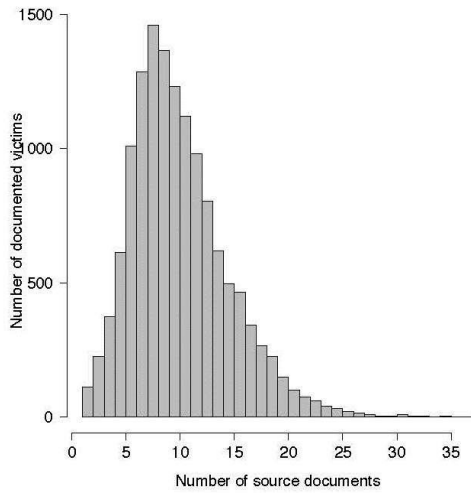
Ball and Kruger Report:

- Based on four visits to HLC
- 64 pages, many tables and graphs
- Reviews database technology
- Descriptive statistics
- Compares HLC to ten other databases

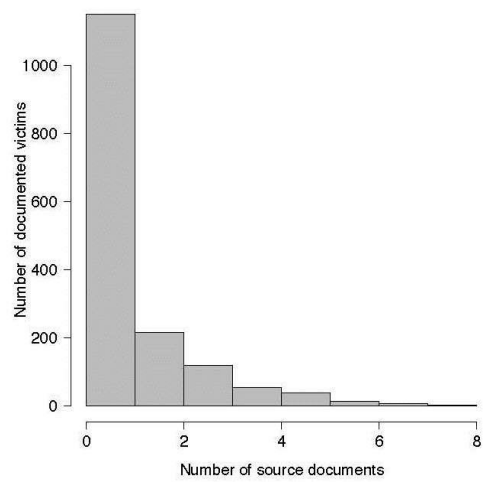
Conclusion: the KMB is very high quality

- Each reported death has many sources
- Each death investigated in detail
- Nearly all victims

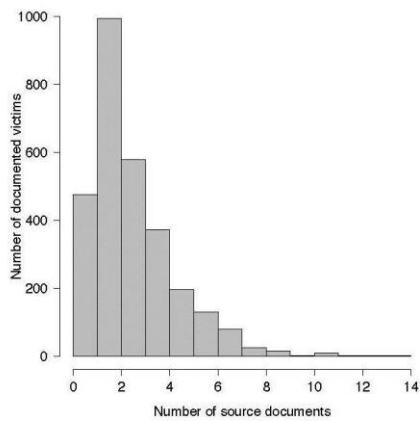
Extensive documentation: source per death



(a) War victims



(b) Potential victims



(c) Not war victims

Figure 5.1: Frequency distribution of source document counts per victim, by victim category.

We conclude that there are very few remaining undocumented war deaths

Table 6.2: Outcomes of three-step record linkage process, and singleton analysis.

Source	A-singletons ^a	HRDAG ^b (Sample) ^c	Matched	(%)	H-singletons I ^d	HLC (Sample)	Matched	H-singletons II ^e
ICRC	63	63 FE	58	92.1	5	5 FE	5	0
OMPF	125	100 RS	53	53	47	10 RS	10	0
OSMN	621	–	–	–	621	60 RS	60	0
MWHK	732	100 RS	64	64	36	10 RS	10	0
LDK	1269	100 RS	59	59	41	10 RS	10	0
ICMP	42	42 FE	39	92.6	3	3 FE	1	2
HRW	125	125 FE	67	53.6	58	10 RS	8	2
ABA	207	207 FE	117	56.5	90	15 RS	8	7
EXH	272	208 AS ^f	146	70.2	62	10 RS	7	3
OSCE	814	200 RS	111	55.5	89	15 RS	12	3

^aSingletons remaining from automated matching.

^bNumber of singletons reviewed by HRDAG.

^cThe type of sample can be a full enumeration (FE), a random sample (RS), or an arbitrary sample (AS).

^dSingletons resulting after HRDAG's first round of hand-matching.

^eSingletons remaining after HLC's second round of hand-matching.

^fA random sample of 100 EXH singletons was drawn, but an additional, arbitrary selection of 108 singleton records was also examined.

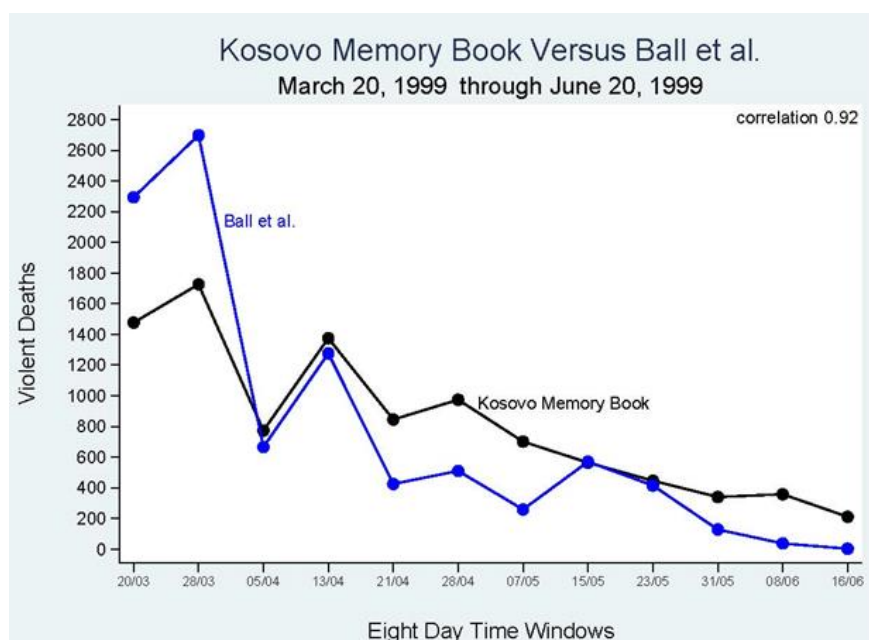
Prof. Michael Spagat:

Okay, so there is an organization known as the Every Casualty Project, and the goal of the Every Casualty Project is to bring about a world where no casualty of armed conflict is left unrecorded. So, from this perspective, this every casualty perspective, the Kosovo Memory Book database is an exemplary achievement, and probably the greatest such achievement in this spirit. Now, what I'm going to talk to you about today actually continues the theme already introduced by Dr. Ball about the virtual comprehensiveness of this Kosovo Memory Book database. We have actually a rather unique opportunity here for assessment and comparison because it turns out that there have been two statistical estimations of the number of people killed in this conflict. Actually, two different estimation projects carried using entirely different methodologies. So, what we have here is actually three separate projects, all measuring the human losses of the Kosovo war. One of them is the Kosovo Memory Book approach of exhaustive documentation, case by case, human being by human being, one by one, okay, each person. The two other projects are using statistical techniques created by statisticians for estimating the total number of deaths without going through with this long exhaustive project, individual by individual, one by one. The first method is, I think, a familiar one to most people – it's the method that's used in political opinion surveys. By this method, what you do is you select, at random, a number of households, in this case it was 1,197 households in Kosovo, you then do interviews at each individual household asking about deaths within that household, okay, and then you measure the percent of people killed within those households that were interviewed, and you extrapolate that in order to create an estimate of the total number of people killed in the entire country. So, if one percent of the people killed within the households that are interviewed, then you estimate that one percent of the entire population was killed. Okay? So, using this household survey method, there is an estimate of 12,000 violent deaths in the Kosovo conflict. That number is actually much closer to the Kosovo Memory Book of 13,535 than it might appear at first glance, because this survey only covers the period from February 1998 to June 1999. Okay? So it doesn't cover the full period. Okay? But, the two sets of numbers are even closer together than that one comparison suggests. Okay? So, here is a month by month comparison. You see the Kosovo Memory Book in black, and the household survey numbers in gold, and it's an extraordinary thing, actually, how closely they track each other. And, you have to bear in mind that monthly estimates using this household survey technique are subject to a lot of random fluctuation, variation, so you would not expect these numbers to match exactly the case by case documented numbers. I also have a report, the full report, that's on the Kosovo Memory Book page, okay, so you can see this picture and other similar comparisons there. Okay? Then, there's another statistical estimate that's done, excuse me, using entirely different methods. It happens that the main person behind this estimate is Patrick Ball, sitting to my left over here, so this technique is a technique that is often used for measuring the size of wildlife populations, it's very different from the household survey technique. The idea of this technique, if you imagine trying to measure the number of fish in the pond, is that you have a net, you dip the net into the pond, you capture a certain number of fish, you place tags on the fish, you mark the fish, put them back into the pond, and then later you come back, you make a second dip of the net, again, count the fish, and importantly you count how many tagged fish come up on the second dip of the net. Intuitively, if a large percentage of the fish on the

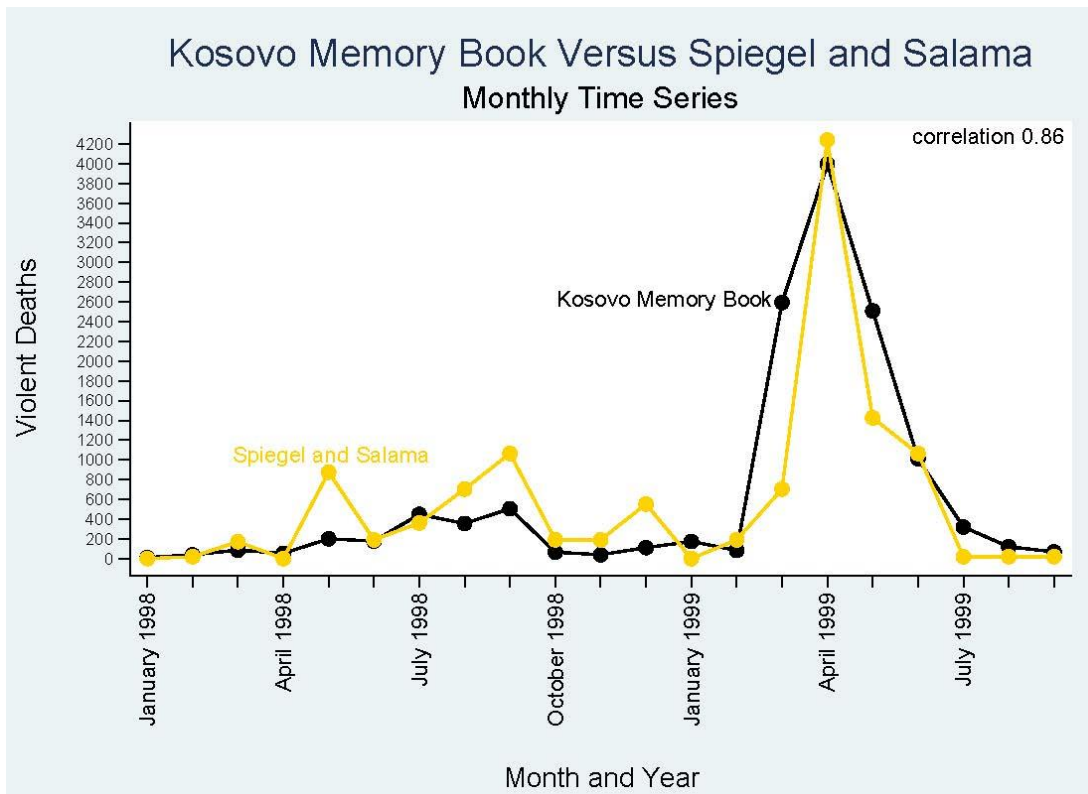
second dip have tags on them, then you think that you've actually captured most fish in the pond both times. If, instead, very few fish come up the second time with tags, than you think that you've actually captured very few of all of the fish in the pond on either one of your dips. There are then statistical techniques proceeding from there to do an estimate of the total number. You, I'm sure, are wondering how that could possibly be relevant here, okay, so the idea is, there are different organizations that are making lists of names of people killed in this conflict. You can match the names across those different lists, and this is analogous to matching fish captured on different dips of the net. Okay? So, that is the idea behind this technique. In the application of this technique, Dr. Ball and colleagues estimated 10,500 deaths in the Kosovo war, and again, that figure is closer to what the Kosovo Memory Book has than it appears at first glance, because this is only covering the period from March 20th to June 12th, 1999. Okay? And, here is a picture, this is not monthly periods, this is actually eight-day periods, it's even harder with eight-day periods to get the statistical estimate to match the pattern of the documentation case by case. So, this is actually a remarkably close match here, okay, given the randomness underlying a sample like this, okay, or an estimate like this. Okay? So, what I am suggesting here is that this, combined with evidence that Dr. Ball showed you a few minutes ago, all points to the conclusion that this is a virtually comprehensive list of every single victim in the conflict. We can never say that we have an entirely complete list, because it's always possible that some further evidence might appear of certain individuals killed, but it seems that, for practical purposes, this is virtually complete. Thank you very much.

The goal of the Every Casualty project is to bring about a “**world where no casualty of armed conflict is left unrecorded.**”
 (<http://www.everycasualty.org/>)

Kosovo Memory Book Database is an exemplary achievement from the point of view of the Every Casualty Project.



Findings The survey included 1197 households comprising 8605 people. From February, 1998, through June, 1999, 67 (64%) of 105 deaths in the sample population were attributed to war-related trauma, corresponding to 12 000 (95% CI 5500–18 300) deaths in the total population.



Through a statistical analysis of these data, this study concludes that approximately 10,500 Kosovar Albanians were killed between March 20 and June 12, 1999, with a 95 percent confidence interval from 7,449 to 13,627. This estimate is consistent with others made by political, legal, and scientific observers.