TOWARDS A RECOMMENDATION FOR A EUROPEAN STANDARD OF PEAK AND LKFS LOUDNESS LEVELS

E.M. Grimm - Utrecht School of Music Technology, Hilversum
R. van Everdingen - Delta-Sigma Audio, Gouda
M. J. L. C. Schöpping - Netherlands Film and Television Academy, Amsterdam

ABSTRACT

Complaints due to loudness variations between program items is a well known issue in television broadcasting world wide. The cause is found in skilled operators being replaced by automated broadcast systems. To facilitate automated leveling of program items, the ITU-R has published recommendation BS-1770 introducing the LKFS loudness measurement. Since the broadcaster's goal is uniform dynamics, a standardized loudness level needs to be accompanied by a standardized peak level. Most stations in Europe use EBU's PPM standard of -9 dBFS and any new standard should comply to this. The analogue PPM falls short in indicating fast (digital) peaks. These do however add to the perceived dynamic range. An additional digital peak standard is therefore needed before an LKFS level can be selected. Our research into the level practice of 50 international television stations suggests setting max digital peak levels to -5 dBFS and nominal loudness levels to -21 LKFS.

INTRODUCTION

The current ITU-R BS.645-2 recommendation for broadcast levels (1) solely specifies PPM (2) peak levels. To harmonize loudness levels between programs, PPM levels fall short. European broadcasters are therefore in need of a new recommendation that includes loudness levels, preferably from the EBU. The new LKFS measurement, introduced in the ITU BS.1770 recommendation (3), is recognized as a very good candidate. To advice on the standard maximum LKFS level, research into the use of dynamic range by 50 television stations was performed. It turned out that an advice on maximum LKFS level should be accompanied by an advice on maximum digital sample peak level.

1. BACKGROUND

Loudness differences between programs and stations are a well known issue (4), (5), (6), (7), (8), (9). The main cause for this is that operators are being phased out in favor of automated play out, human level control is lacking in modern facilities. Another cause in The Netherlands is that Dutch stations have adopted the policy to transmit material without alteration. Reason is that in 1999 an advertiser consulted an engineering agency to perform level measurements on commercial blocks. It turned out there were differences in peak (PPM) levels (10). The advertiser threatened to withdraw if the station would change their commercials level again. The station was only allowed to check compliance to the ITU/EBU recommendation, which solely addresses max PPM levels (1). Loudness is determined by average energy, not by peaks, so loudness can vary considerably between programs with equal peak levels. Result of the new policy was that viewers started complaining about the

commercials being too loud. Since riding levels was not an option anymore, the only way out was to raise the level of normal program material by means of general dynamic range compression. This appeared to solve the problem just partly because programs like a documentary with dialog recorded on a market square loose intelligibility as a result of compression much sooner than close miked commercial voice overs. A compromised setting of the compressor was found. Viewer complaints were reduced, but did not vanish. The new setting however did lead to complaints by directors and engineers of drama and documentary films because it had negative effect on their artistic mix balance choices and caused loss of dramatic impact by compromising deliberate level differences.

In May 2005 the 'VCA', a society of Dutch Film Sound Engineers (11), organized a "Television Broadcast Loudness Symposium" (12) in The Netherlands that was visited by representatives of all relevant fields: engineers and managers of public and commercial broadcast stations, post production houses and rebroadcasters and also producers and directors of commercials, drama and documentaries. In the end everybody agreed on the fact that the current situation was unfavorable. Even producers of commercials agreed they were longing for the artistic options of dynamic range in sound like in drama productions. Currently they are forced to use heavy compression because otherwise their clip will be too low against the competition in a commercial block. A different kind of problem was posed by the rebroadcasters that have to combine the levels of several stations in such a manner that viewers can switch between channels without large level differences. They had started developing automated systems for gradually leveling loudness differences and longed for a standard in loudness measurement.

It was decided to create a Dutch Broadcasting Loudness Committee on which the authors of this paper sit, along with representatives of public and commercial stations and post production houses. Their task would be to research a means for leveling program material while preserving artistic choices. An addition to the current audio levels standard - incorporating a loudness level measurement - was soon found to be necessary. International harmonization for this measurement was seen as mandatory and cooperation with ITU's SRG3 group started. Their BS.1770 LKFS loudness measurement (3) appeared to be perfectly suitable for Dutch program material and lead to better consistency in its program loudness levels than overall dynamic range processing like in use at the moment (13). Other authors report similar results for other countries (14).

1.1 Dynamic range

When it came to implementing BS.1770, the committee soon faced a complicated issue. Setting a loudness level only makes sense in its relation to a peak level. There is demand for both uniform loudness levels and enough dynamic range. Dynamic range can be expressed by the difference between average loudness level and the maximum digital (sample) peak level. Different environments demand for different dynamic ranges. The dynamic range preferred in a movie theatre is about 6 to 7 dB larger than the optimum for a domestic environment (15), (16). This means that once the maximum digital peak level is set, the loudness level can be chosen accordingly.

The problem we encountered was that although PPM levels are standardized in the EU, digital peak levels are not. EBU recommends 0 PPM to be aligned to -9 dBFS (17), the PPM meter having attack ballistics of 10 ms (2). But on digital peak levels (17) states that "due to the characteristics of quasi-peak program meters used by broadcasters, the true program peaks can be 3 dB greater than those indicated; When operator errors are taken into account the true peaks may occasionally be 6 dB greater than indicated". Digital peaks above -9 dBFS are mentioned, but there is no official limit to the maximum digital peak level in this recommendation.

Current practice in many broadcast stations has become the use of very fast look ahead 'brick wall' limiters. In the days of analog on air transmission these were used to limit the deviation and corresponding bandwidth of the transmitter. The lack of an indication for maximum levels in the EBU recommendation (17) has probably lead to confusion about the level where this brick wall limiter should start to act. This in spite of the PAL B/G standard of ITU-R (18) and its European adaptations like CENELEC EN50083-5 (19). These clearly advice to align 0 PPM to 30 kHz FM deviation, which is 4.4 dB below the allowed maximum peak level. Some stations (like the national German stations) followed the standard and had their limiters set accordingly to about -4.5 dBFS. Some have interpreted the EBU recommendation (17) and ITU-R recommendation (1) such that "no signal may exceed -9 dBFS" (20) and set their limiters to -9 dBFS. Since operators often monitor on PPM's referring to the ITU-R recommendation (1), this last situation unfortunately leads to quite severe peak limiting.

Lately analog on air transmission of television stations is being phased out. This creates a situation where analog RF modulation is only performed at rebroadcaster facilities for cable television. Here dozens of local, national and international channels, with audio sources varying from terrestrial and satellite reception to direct studio links are prepared for delivery. And this preferably such that the viewer does not experience large level jumps when switching channels. Since most rebroadcaster do not want to interfere with the dynamic range chosen by stations, the loudness level of the channel with least dynamic range is attenuated to match that of the channel with largest dynamic range. Modern technology like LKFS based automatic level control systems (21) are installed in some cable head-ends and recently even audio level control directly in MPEG streams (without de- and recoding) was introduced to facilitate leveling in DVB links (22). Of course these systems cannot cater for level differences within one channel, they only minimize level changes between channels.

Consequence for broadcasting stations is that choices in types of processing in the transmission chain will not influence the relative loudness of the station on the viewer's television (like it did in the on air transmission days). It will only change the available dynamic range for program material and thus the sound quality. Stations with smaller dynamic range just end up peaking at lower levels in the viewer's television than stations with higher dynamic range. This is confirmed by measurement of the authors. Very fast peak limiting to avoid overload of analog RF modulators therefore makes no sense at all when combined with dynamic range compression.

2. SURVEY OF CURRENT LEVELS

As mentioned before: if we would like to solve the annoying level differences between programs while preserving the director's artistic choices in use of dynamic range, an LKFS based maximum loudness level seems like the appropriate choice. However, before we can select a standard LKFS level for EU broadcast stations first consensus is needed about the maximum allowed digital peak level. We strongly recommend keeping the old ITU-R recommendation (1) but to add a recommendation for both maximum digital peak level and maximum LKFS level to it.

We decided to perform a survey on current digital peak, PPM and LKFS levels in Europe. Based on this knowledge a proper choice can be made selecting the right maximum levels for digital peaks and LKFS. In our opinion any new standard recommendation needs to comply with daily practice as seamless as possible.

2.1 Survey design

To have an objective view on current practice one needs access to raw transmissions of many stations. Reception from analog cable delivery is ruled out due to interference of the

rebroadcaster with levels. Digital television reception showed to be the best compromise. Although some stations are known to process their DVB feeds differently from their direct feeds to analog transmitters and rebroadcasters, on average it provides a representative view. We selected 50 international stations, listed in table 1. The channels were received through Dutch rebroadcasters REKAM and CAIW. These companies mostly use satellite links for reception of foreign stations, the latter however also uses direct studio feeds for national and local stations. REKAM provides a level transparent link, CAIW utilizes the new ITNM Systems ALC system to losslessly control gain directly in MPEG streams (22). The ALC levels were known to us, so we could correct the measurements correspondingly. 46% of the stations was Dutch, due to the goals of the Dutch Broadcasting Loudness Committee. This still leaves 27 channels of foreign origin, with 24 coming from EU stations (apart from the Dutch). Reasonably representative therefore for the European situation.

In this paper we have chosen to not anonimize the names of the stations. Reason is that we have reached a stadium of discussion about practical implementation of LKFS and Peak level standards. We feel it is important to have a view on everyone's position in the debate. By no means we have the intention to reveal 'bad' behavior. Every station has had its valid professional grounds to set the levels as they did.

We have chosen to not make a very detailed analysis of just a few channels, but a broader analysis of many channels instead. This provided us with a wider view that seemed advantageous to us. We recorded a few different programs on several days for every channel, each fragment being about 6 minutes in length. The digital peaks were measured using the internal meters of the Adobe Audition software used for recording (these are sample peak levels, not true peak levels as advised in ITU-R recommendation BS.1771 (23)). PPM levels were measured using a calibrated DK Audio MSD100. LKFS levels were measured using a dedicated piece of software, developed at Utrecht School of Music Technology by Wouter Snel ea (13). This software fully complies to BS.1770.

3. RESULTS AND ANALYSIS

In graph 1 the levels of all measurements are gathered. These are the exact levels received, without any normalization. Maximum LKFS is plot in blue, maximum PPM in yellow and maximum Digital Peak in red. Investigating the LKFS levels in graph 1 shows very large loudness differences among channels. Remind these are the levels as transmitted by the stations, without rebroadcaster interference. On analog cable reception the level differences are potentially smaller because of rebroadcasters controlling levels. Our measurements show the level selection of the stations themselves, as received by rebroadcasters. The most extreme loudness difference is between Het Gesprek (-29 LKFS) and NET5 (-13 LKFS): a jump of 16 dB in loudness! One step better still shows 9 dB in loudness difference: BBC2 (-24 LKFS) versus SBS6 (-15 dBFS). It is clear that if stations do not harmonize their levels, rebroadcasters will have to do it for them.

As mentioned in paragraph 1.1, some stations have chosen to set their brick wall limiters to a threshold of -9 dBFS. These can easily be tracked in graph 1, about 22% of the measured stations limit to -9 dBFS +/- 1 dB.

Another interesting element found in graph 1 is that many stations do not comply with ITU-R BS.645-2 (1). Only 40% of the stations peak within 1 dB of 0 PPM. 34% peak too high, 22% peak too low. We kept two stations out of these percentages because of inconsistency in their broadcasts: sometimes peaking 9 dB low, sometimes 3 dB loud... The EBU recommendation is followed by less than half the stations, which we found quite amazing.

3.1 Normalized to 0 PPM

We decided to virtually "correct" this omission by normalizing the graph to have all stations peak at 0 PPM = -9 dBFS. This is shown in graph 2. Here can be seen what happens if the current ITU-R recommendation (1) would be observed by everybody, without changing other routines. Some stations show no digital peaks above 0 PPM (peaking at -9 dBFS), others peak higher, even up to +1 dBFS! About 37% of the measured channels have less than 2 dB difference between 0 PPM and maximum digital peak. The majority keeps more headroom above 0 PPM, although many do not seem to align 0 PPM to -9 dBFS.

Watching the LKFS levels in graph 2 is an educational experience. Loudness levels still vary widely. From -28 LKFS (TRT International) up to -16 LKFS (Nederland 2, Nederland 3, RTL Television): a difference of 12 dB. A peak standard like ITU-R BS.645-2 (1) clearly does not regulate loudness. Loudness level differences can also be found between different measurements of the same station. Some of these are due to fragments in- or excluding commercial breaks, although in our collection this only brought about 2 dB differences. Comparing several 6 minutes program fragments of BBC1 however showed averaged loudness levels ranging from -23 LKFS to -16 LKFS. Level consistency currently can be pretty low when normalizing 0 PPM peaks.

3.2 Normalized to LKFS

What matters to viewers is harmonized loudness levels, the LKFS value. A new recommendation has to include this. But what level should it be? This is a question about desired dynamic range ("DR"). Since DR is defined by the difference between perceived loudness level (LKFS) and digital peaks, the first target is to determine the maximum digital peak level. Our opinion is a new recommendation should not interfere with stations featuring a reasonably wide DR. When low DR stations line up LKFS level with them, no one has to change its routine. The low DR stations just don't peak to max anymore. The opposite choice - to have wide DR stations line up LKFS with low DR stations - leads to wide DR stations losing their character due to peak limiting. Now let's have a look at current use of dynamic range in our collection of stations.

19 to 20 dB: BBC1, Canvas, TRT International, Omroep Zeeland

17 to 18 dB: France2, TV Polonia

15 to 16 dB: BBC2, Animal Planet, National Geographic, VRT1, TV5, WDR3, TVE, RTV Noord, RTV Drenthe, RTV Oost, TV West, TV Rijnmond, TV Utrecht, Ketnet, Het Gesprek, Chinese News & Entertainment.

13 to 14 dB: Nederland 2 (movie), Net 5, BBC World, SWR, ZDF, Euronews, Eurosport, Discovery, RTL 8, RTL Television, Rai Uno, Omroep Fryslan, Omroep Gelderland, Omroep Brabant, L1 Limburg, AT5.

11 to 12 dB: Nederland 3 (movie), SBS6, ARD, RTL4, 5, 7, La Une, CNN, Nickelodeon, Omroep Flevoland.

9 to 10 dB: Nederland 1, 2, 3, RTV Noord-Holland.

Taken into account our premise that a new recommendation should not interfere with wide DR stations, the DR must be 16 dB. DR's of 17 to 20 dB were only caused by occasional peaks. Such a wide DR is also objectionable in a domestic environment.

Next question is what maximum digital peak levels do we find? If we analyze the maximum digital peak levels of the "15 to 16 dB" DR group in graph 2 we get results ranging from -7 dBFS to -2 dBFS. 56% had peaks between -5 dBFS and -6 dBFS. Higher peaks were quite rare, stations with occasional -2 dB peaks more often peaked at -6 dBFS. It seems safe to

select a maximum digital peak level of -5 dBFS, this is 4 dB above the 0 PPM alignment level.[1] Subtracting a dynamic range of 16 dB from -5 dBFS leads to a maximum loudness level of -21 LKFS. Incidently the -5 dBFS peak level corresponds with the PAL B/G recommendation (18).

In graph 3 the loudness level of all fragments is normalized to -21 LKFS. When set like this, switches between programs and stations will only show minor loudness differences. As can be expected peak levels in this graph vary widely since current dynamic range of stations is divergent. Peak levels may vary in the future too, due to artistic choices by directors.

4. CONCLUSION AND PROPOSAL

We have investigated the use of dynamic range of 50 television stations by measuring their LKFS, PPM and digital peak levels. We advice a new audio level recommendation to include LKFS loudness levels and digital peak levels. We advice it should still comply to the current ITU-R/EBU recommendation for max PPM levels (1). We advice it should not limit the dynamic range demand of the majority of stations. We found a dynamic range between LKFS and max digital peak of 16 dB to fit that requirement. Stations using this dynamic range and peaking at 0 dB PPM showed digital peaks up to about -5dBFS. Our proposal therefore is to set the maximum digital peak level to -5 dBFS and the maximum loudness level to -21 LKFS.

A recommendation like this could well be communicated to all suppliers of content. Necessary measurement equipment is readily available. Anyone who for artistic reasons does not wish to use the full dynamic range bandwidth will automatically be aligned to -21 LKFS level and thus be equally loud to programs with wider dynamic range. He is just peaking at lower levels. On the other hand directors that do wish to tell their story using loudness contrasts will have proper freedom to do so, be it in a documentary or a commercial. Rebroadcasters can use the same recommendation, in fact some of them comply already by using LKFS in their automated leveling systems and leaving enough headroom for stations with high dynamic range demands.

4.1 Control

No audio level recommendation can be successful without control tools. We have the following suggestions for the various stages in the broadcast chain. Further research into these is recommended

a. In Digital Media Asset Management Systems, automatic program level attenuation down to the standardised LKFS level can be performed at the moment of intake. Implementation of this function has to be be investigated.

b. -21 LKFS is the maximum loudness level, but programs with too low loudness levels are also objectionable for the viewer. Automatic amplification at the intake stage of (a) might be an option here. We advice to set an LKFS threshold for amplification lower than -21 LKFS to cater for 'film noir' like movies being softer on average. -26 LKFS would be an educated guess, but research is needed to find the right amplification threshold level.

c. Any system should be able to deal with dynamic range accidents like often encountered in live broadcasts or when the cinema mix edition of a movie has been loaded in the Media System. In cooperation with Orban Europe we have performed an experiment where an audio processor was set to simultaneously have an instantaneous limiter for digital peaks

---

[1] In case of analog transmission feeds pre-emphasis limiters can be set to the same -5 dBFS (24).

above -5 dBFS and a 10 ms attack time limiter for peaks above -9 dBFS (mimicking a PPM). This appeared to impose the peak recommendation perfectly: PPM's never crossed 0 PPM, digital peaks never exceeded -5 dBFS. It sounded good too.

5. REFERENCES

1. ITU-R Recommendation BS.645-2, 1992. Test signals and metering to be used on international sound-programme connections.

2. IEC Standard 60268-10, 1991. Sound system equipment - Peak programme level meters.

3. ITU-R Recommendation BS.1770, 2006. Algorithms to measure audio programme loudness and true-peak level.

4. Klar, S. and Spikofski, G., 2002. On levelling and loudness problems at television and radio broadcast studios. Audio Engineering Society Paper 5538 presented at the 112th Convention.

5. Emmett, J., 2003. Audio levels — in the new world of digital systems. EBU technical review Januari 2003.

6. More, B. C. J., Glasberg, B. R. and Stone, M.A. 2003. Why are commercials so loud? - Perception and Modeling of the Loudness of Amplitude-Compressed Speech. J.Audio Eng.Soc., Vol.51, No.12, 2003 December

7. Klar, S. and Spikofski, G., 2004. Levelling and loudness in radio and television broadcasting. EBU technical review January 2004.

8. Moerman, J.P., 2004. Loudness in TV-sound. Audio Engineering Society Paper 6039 presented at the 116th Convention.

9. Spikofski, G., 2006. Who can do louder? Loudness jumps in broadcasting and how to avoid them? 24th Tonmeistertagung International Convention November 2006.

10. De Leeuw, D. 2005. Luidheidniveaus Reclameblokken; gemeten in de jaren 1999 - 2002. Adviesbureau Peutz. Presentation at Luidheidsymposium Televisiegeluid May 2005.

11. Vereniging Constructief Audio. http://www.vcafilmsound.nl

12. Luidheidsymposium Televisiegeluid, 't Spant Bussum, May 14th 2005.

13. Snel, W., 2006. Luidheid in de praktijk - een onderzoek naar de praktische invoering van luidheidscontrole voor de Nederlandse televisie. USMT master thesis.

14. Lund, T, et al, 2008. 124th AES workshop W7. Loudness - a change of paradigms?

15. Dolby Laboratories, 2000. Dolby Digital Professional Coding Guidelines. p88.

16. Katz, B. 2000. Integrated Approach to Metering, Monitoring, and Leveling Practices, Part 1: Two-Channel Metering. J.Audio Eng.Soc., Vol.48, No.12, 2000 December.

17. EBU Recommendation R68-2000, 2000. Alignment level in digital audio production equipment and in digital audio recorders.

18. ITU-R Recommendation BS.707-5, 2005. Transmission of multisound in terrestrial television systems PAL B, B1, D1, G, H and I, and SECAM D, K, K1 and L.

19. CENELEC EN 50083-5, 2003. Cable Networks for Television Signals, Sound Signals and Interactive Services - Part 5: Headend Equipment.

20. NOB, 2005. SLA-E1 - Eisen aan informatie en materiaal (E1).

21. Jünger Audio C8000 Level Magic audio processors.

22. ITNM Systems ALC100 DVB Audio Level Control.

23. ITU-R Recommendation BS.1771, 2006. Requirements for loudness and true-peak indicating meters.

24: ITU-R BS.642-1, 1990. Limiters for High-Quality Sound-Programme Signals - Section 10B - Frequency-Modulation Sound Broadcasting in Bands 8 (VHF) and 9 (UHF).

ACKNOWLEDGEMENTS

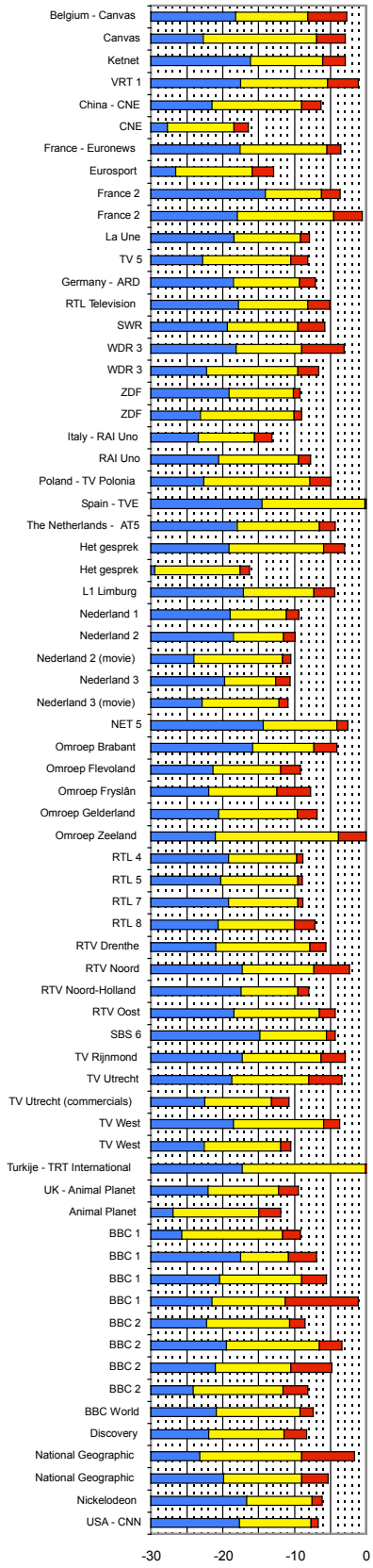
The authors would like to thank their collegues of the Dutch Broadcast Loudness Committee for all our valuable discussions, Wouter Snel for his cooperation in this project and the VCA for starting it all, back in 2004.

Table 1. List of measured stations.

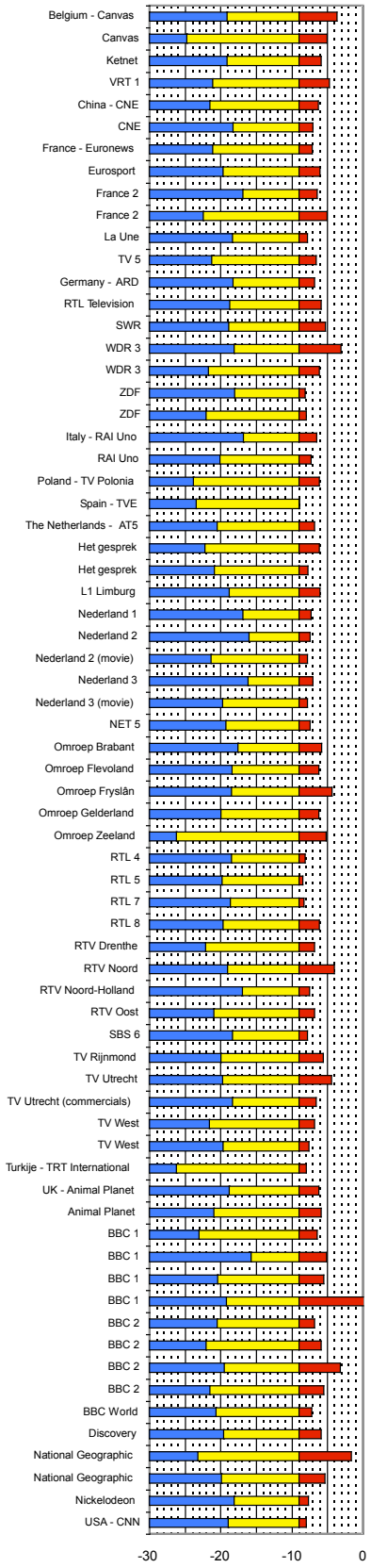
Belgium - Canvas, Ketnet, VRT 1. China - Chinese News & Entertainment. France - Euronews, Eurosport, France 2, La Une, TV 5. Germany - ARD, RTL Television, SWR, WDR 3, ZDF. Italy - RAI Uno. Poland - TV Polonia. Spain – TVE. The Netherlands - AT5, Het gesprek, L1 Limburg, Nederland 1, Nederland 2, Nederland 3, NET 5, Omroep Brabant, Omroep Flevoland, Omroep Fryslan, Omroep Gelderland, Omroep Zeeland, RTL 4, RTL 5, RTL 7, RTL 8, RTV Drenthe, RTV Noord, RTV Noord-Holland, RTV Oost, SBS 6, TV Rijnmond, TV Utrecht, TV West. Turkije - TRT International. United Kingdom - Animal Planet, BBC 1, BBC 2, BBC World, Discovery, National Geographic Channel, Nickelodeon. USA – CNN

Notes to graph 1, 2 and 3.

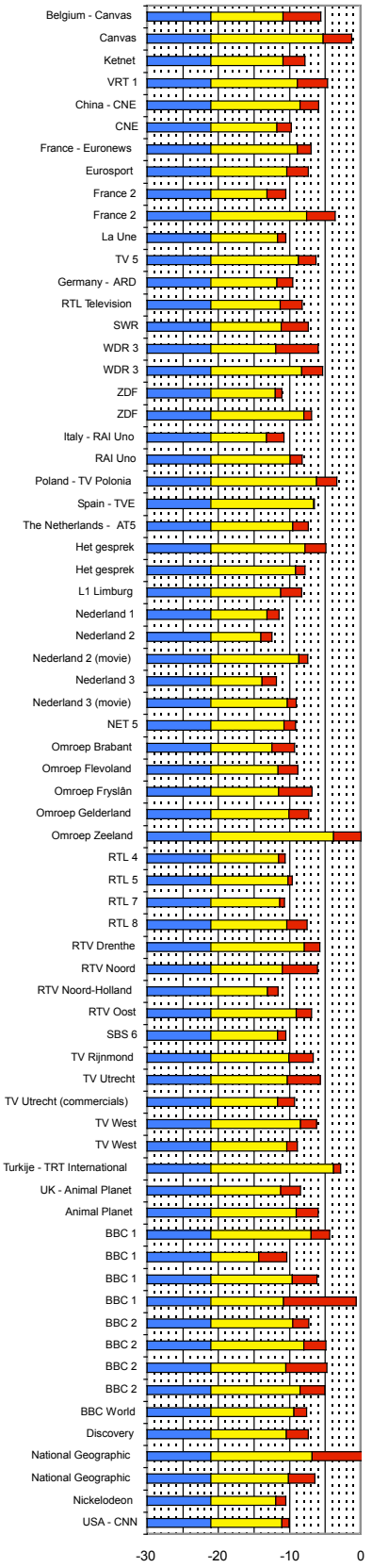
Measurements of all table 1 stations are presented. Blue level is max LKFS, yellow level is max PPM, red level is max digital peak. Channels presented with 1 bar were consistent. The LKFS levels of all measured fragments have been averaged, the max PPM and max Digital Peak levels show the highest value. This is comparable to a long measurement, overlapping the content of the seperate measurements. If a channel is present with more then 1 bar, the consistency of (some of the) fragments was low because differences in level were larger than 3 dB. Averaging would not be representative.



Graph 1.
Measurements of 50 stations: LKFS loudness, PPM peak and digital peak.



Graph 2.
The same measurements, normalized to all ppm max levels equal -9 dBFS.



Graph 3.
The same measurements, normalized to all LKFS levels equal -21 dBFS.