ICPAC QUALITY CONTROL ASSURANCE AND PROCEDURES

It is widely known that raw digital climate data are subject to a wide range of errors, which can be introduced along the chain of collecting, digitizing, processing, transferring, storing and transmitting the data (Anonymous 2010). A wide set of procedures for quality controlling (QC) raw data have been developed, which are recommended by the World Meteorological Organization (WMO) and implemented or used by investigators to detect and flag potentially erroneous values in any dataset (e.g. Abbott 2sgma check1986).
The 2sgma check and neighboring station check has been done to the data. The Consistence check to compare the available records from different sources has also been made. ICPAC, as an RCC is obliged to support NMHS build capacity in climate data management which includes climate data quality control.
As mentioned before ICPAC has no mandate to observe climate data, rather receives digital data from 137 IGAD NMHS station data on bilateral agreements on a 10day basis. This is augmented with the rainfall estimates and other datasets from the public domain like http://climexp.knmi.nl/selectdailyseries.cgi which serves in climate monitoring. There are two levels of quality control at ICPAC and they are Primary and Secondary level of quality control.

Primary level QC

Observational data error detection and possible error corrections in order to ensure the highest possible reasonable standard accuracy for real-time can follow the following steps;

I. Quick inspection of the data to isolate anomalous deviations and suspicions zeros in a space of missing data.
II. Action taken in the real time situation is to contact the NMHS where data originated by phone or email such that the person in charge quickly confirms the validity of the value.
III. If the value is too suspicious beyond 4sgma is put aside for further investigation.

Secondary level QC

When data is received at ICPAC it is subjected to further rigorous checks to come up with quality controlled climate data series for long range monitoring, detection and attribution. Following are some of the checks employed using both ***Macros*** and ***Climsoft***:

- Repeated data check
- Consistent checks with other data existing at ICPAC
- 2sgma Check for Dekadal monthly data sets
- 10th Percentile check for the Rainfall Lower limits
- Daily rainfall checked with upper limit of 100mm but this can change according to location
- Temperature rates of change from the previous day

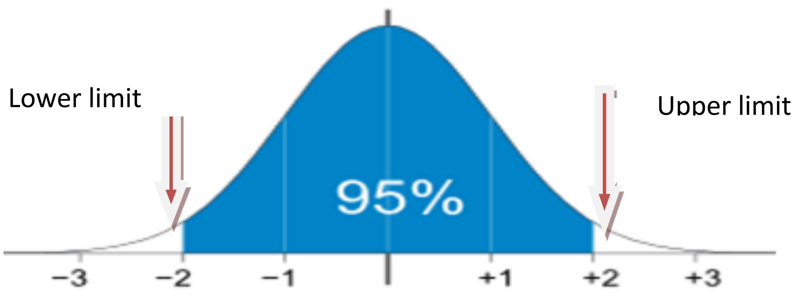
The 2sgm Test:

The distribution of rainfall and temperature follows 2sgma away from the mean as recommended by Abbot in 1985 and can be used to set lower and upper limits checks. Both errors and extreme values sit in the tails beyond the lower and upper limit but this can only be true if the data is having a normal distribution. In the case of the skewed data like rainfall, the upper limit is used and the lower limit is set at the 10th percentile .However in the case of temperature which is normally distributed, this formulae works for the lower and upper limit

Organize the data you need to quality control whether Monthly or dekadal data from which the statistics can be generated.

Getting the upper limit to be used to quality controlling the rainfall data using the 2sgma and this is before you use the macro.

$$\textit{Upper and lower limits} = \bar{X} \pm 2\partial$$



Performing Quality Control using Climsoft

Data brought in to ICPAC from the regional NMHS is already digitized and can contain inherent errors due observation or transmission which may have not been detected from the source. The data is imported into Climsoft database and subjected to further quality control checks.

Outliers causing noise in the data and Inconsistencies between data items for example, the minimum temperature <= maximum temperature for the same period at the same location can be detected by Climsoft which has inbuilt modules; these checks can be performed as instructed below:

Before performing any checks, it is important to establish the plausible upper and lower limit of the elements you are to check; some people prefer to put the absolute or highest value ever recorded. This is not advisable because the return period of such values can be more than 30years. The lower and upper limits for precipitation and temperature generated at ICPAC for each stations and then grouped as par climatic zones for easy of operations, The table below shows these grouped limits.

Areas of interest	Rainfall		Tmax		Tmin	
Limits	Lower	Upper	Lower	Upper	Lower	Upper
Cold places	0	100mm	7°C	24°C	-2°C	16.5°C
Places with winter	-	-	0°C	20°C	-4°C	10°C
Coastal regions	0	150mm	15°C	34°C	14C	24°C
Mountainous regions	0	130mm	8°C	26°C	-2.1°	16°
Lift valleys	0	70mm	16°C	35°C	14.5°	23°
Deserts& marginal areas	0	60mm	20°C	38°C	13.6°C	25°C
Warm Places	0	100	13.5°	33.5°C	10°C	22°C

Setting the limits

When logged on Climsoft the welcome data management form will appear as shown in the Fig1a then click on the Metadata icon and in then in check on the element Fig1b. This will give you an interface to set the element upper and lower limits as shown in Fig2a, then click on update to save the changes.

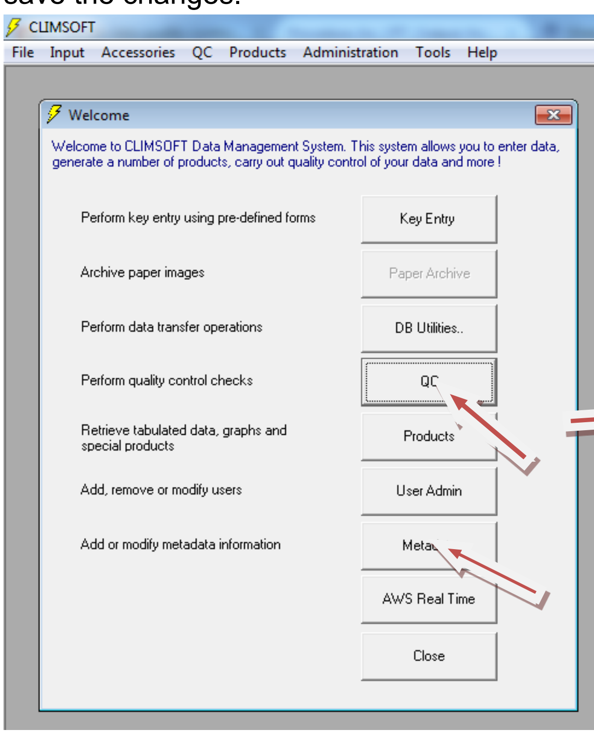


Fig1a

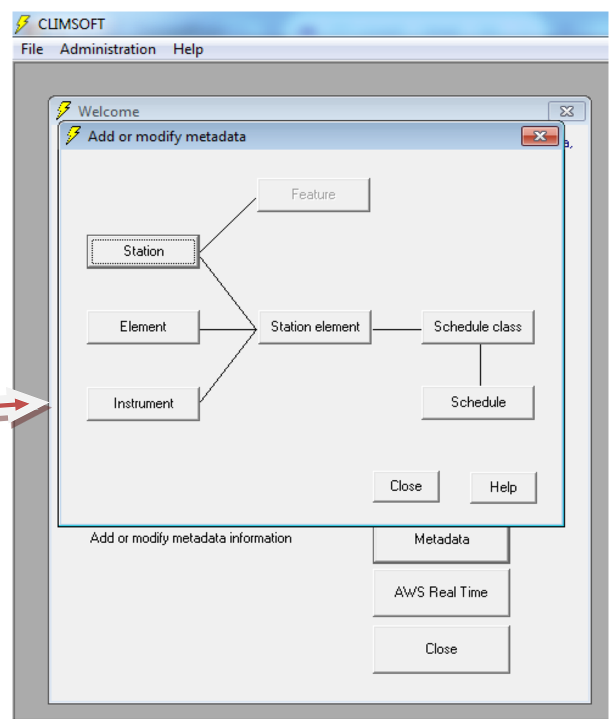


Fig1b

After setting the limits and still logged on Climsoft on the welcome data management form Fig1a click on the QC icon to begin the quality control process, then check in the absolute limit as shown in Fig2b below and click ok

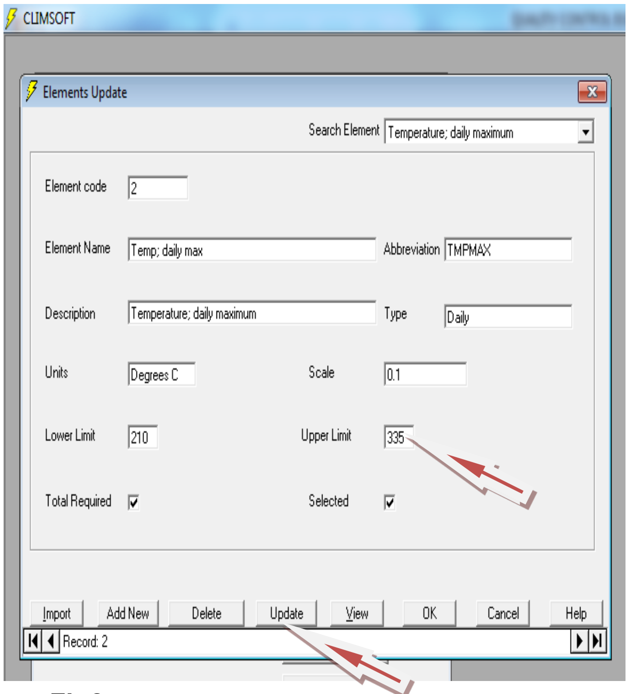


Fig2a

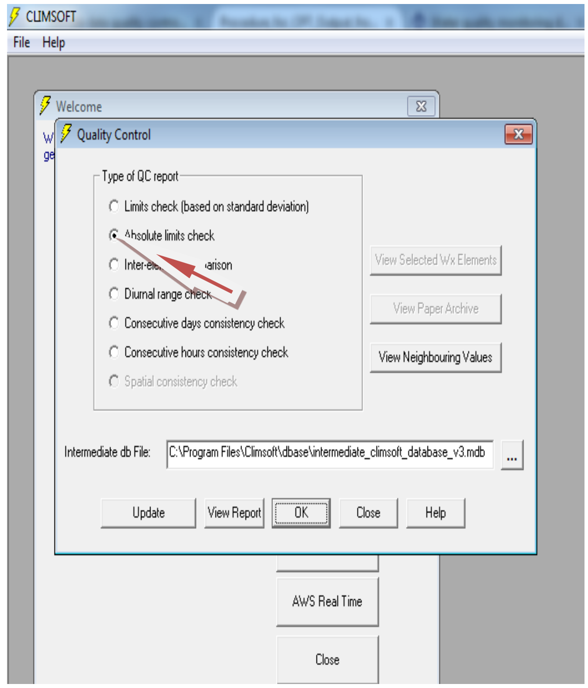


Fig2b

The output excel table containing the values to be checked for their correctness will appear as below. NB the output values have to be scaled down because they were factored by 10.

OUTPUT TABLE FOR THE STATION WITH VALUES EXCEEDING THE LIMITS

station	elemer	obs_datetime	yyyy	mm	dd	obs_leve	obs_value	limit	val_limit_	limit_type	qc_stat	acquisitic
74360150	5	18/08/1964 06:(	1964	8	18	surface	809	700	-109	Upper_limit	1	1
74360150	5	18/07/1973 06:(	1973	7	18	surface	703	700	-3	Upper_limit	1	1
98330160	5	30/11/1997 06:(	1997	11	30	surface	876	700	-176	Upper_limit	1	3
98330160	5	19/02/1998 06:(	1998	2	19	surface	836	700	-136	Upper_limit	1	3
74360150	5	15/07/1974 06:(	1974	7	15	surface	865	700	-165	Upper_limit	1	1
74360150	5	24/08/1987 06:(	1987	8	24	surface	900	700	-200	Upper_limit	1	1
74360150	5	25/06/1996 06:(	1996	6	25	surface	734	700	-34	Upper_limit	1	1
95320520	5	18/12/1979 06:(	1979	12	18	surface	721	700	-21	Upper_limit	1	3
95320520	5	14/05/1982 06:(	1982	5	14	surface	792	700	-92	Upper_limit	1	3
95320520	5	19/04/1985 06:(	1985	4	19	surface	735	700	-35	Upper_limit	1	3
95320520	5	20/03/1988 06:(	1988	3	20	surface	718	700	-18	Upper_limit	1	3
95320520	5	19/03/1990 06:(	1990	3	19	surface	764	700	-64	Upper_limit	1	3
10350260	5	24/12/1974 06:(	1974	12	24	surface	728	700	-28	Upper_limit	1	3
95320520	5	18/03/1995 06:(	1995	3	18	surface	1555	700	-855	Upper_limit	1	3
95320520	5	28/11/1996 06:(	1996	11	28	surface	760	700	-60	Upper_limit	1	3
95320520	5	28/12/1996 06:(	1996	12	28	surface	760	700	-60	Upper_limit	1	3
95320520	5	18/04/1997 06:(	1997	4	18	surface	814	700	-114	Upper_limit	1	3
10350260	5	13/01/2002 06:(	2002	1	13	surface	704	700	-4	Upper_limit	1	3
10350260	5	21/02/1978 06:(	1978	2	21	surface	885	700	-185	Upper_limit	1	3
10350260	5	21/12/1982 06:(	1982	12	21	surface	764	700	-64	Upper_limit	1	3
10350260	5	18/01/1986 06:(	1986	1	18	surface	856	700	-156	Upper_limit	1	3
10350260	5	21/01/1986 06:(	1986	1	21	surface	896	700	-196	Upper_limit	1	3
10350260	5	19/03/1986 06:(	1986	3	19	surface	777	700	-77	Upper_limit	1	3
10350260	5	28/03/1989 06:(	1989	3	28	surface	718	700	-18	Upper_limit	1	3
10350260	5	29/01/1991 06:(	1991	1	29	surface	740	700	-40	Upper_limit	1	3

Check the output data against the original to see if it is not a mismatch. From above table, the station within this region falls under this set limit of 700mm except Station ID No 95320520 exceeding the limit with 85.5 mm is to be checked further with neighboring stations to confirm its validity. NB It require a qualified and experienced climatologist to do this kind of work not to discard true anomalous values. It is recommended to flag these values and keep the database as original as possible

Steps to run the quality Control Macro

These ICPAC Macros were developed to do further checks like repeated data, temperature step checks, temporal and spatial consistence of a station, They are used at post data ingestion level away from the DBMS and they are very easy to use. They are intended to manifest the errors or suspicious values which were undetected at the lower level base level before the data can be used for analysis to generate dependable results. The steps used to run the macros are listed below.

a) Open the excel file containing the Macro;
b) Enable the macro by clicking option and enable Macro before it can run;
c) Select and Copy the original dataset and paste it in sheet1 of the macro;
d) Ensure that the data starts on column "I" for the Monthly data;
e) Remove all the -999 from the dataset and replace them with blank;

f) Pre-calculations of Upper and Lower limits have been embedded within sheet macro as it can be seen in the table below;
g) Press Cntr_Shift_U for the Upper-limit to run which will highlight or flagged values exceeding the limit;
h) Press Cntr_Shift_D this runs temperature rates of change
i) Press Cntr_Shift_L to run the Lower-limit separately in case ;
j) Press Cntr_Shift_R to run the Repeat check macro;
k) Press Cntr_Shift_C to clear the flags ;

After running the macro the anomalous values are flagged in blue or red depending on the magnitude of the error.

Step2

This is investigation of the validity of the suspicious values; it should be noted that not all flagged values are wrong. For example the blue anomalous value in Jan 1998 was due to El-Nino and then the red value of 677 in Sept 1986 was missing a decimal point. This kind of work requires patience and a critical mind with lots of experience in climatology.

MACRO EXCEL TABLE WITH FLAGGED VALUES

Country	Station	Source	Latitud	Longitu	Elevatio	ID	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Kenya	MWIRIM	MET	45	60.62	515	91340240	1981	85.3	98.7	45.1	165.1	221	235.1	78.4	168.3	78.9	39.7	36.2	28.9
Kenya	MWIRIM	MET	45	60.62	515	91340240	1982	9.7	0.4	32.9	114.4	204.8	138.4	128.3	111.2	92.1	29.9	65.2	21.2
Kenya	MWIRIM	MET	45	60.62	515	91340240	1983	3.7	11.2	341	274.7	174	67.6	141.2	171.6	240.9	120.1	19.5	51.8
Kenya	MWIRIM	MET	45	60.62	515	91340240	1984	6.4	48.1	54.3	253.1	206.7	78.3	151.5	199.7	62.7	158.5	274.6	43.3
Kenya	MWIRIM	MET	45	60.62	515	91340240	1985	26.3	107.8	24.7	100.6	188.3	160.9	109.8	182.3	173	216.1	119.5	36.3
Kenya	MWIRIM	MET	45	60.62	515	91340240	1986	9.3	10.2	23.6	102.2	138	122.7	121.1	114	677	46.2	106	52.7
Kenya	MWIRIM	MET	45	60.62	515	91340240	1987	26.8	35.9	190.8	221.2	213.1	68.2	118	156.3	62.2	24.1	99.8	2.9
Kenya	MWIRIM	MET	45	60.62	515	91340240	1988	3.1	16.6	87.5	213.5	325	69.7	184.8	114.1	47.2	73.6	31.3	21.2
Kenya	MWIRIM	MET	45	60.62	515	91340240	1989	12.2	79	118.9	179.6	236.9	66.8	32.9	184.7	43.4	61.2	135.4	21.3
Kenya	MWIRIM	MET	45	60.62	515	91340240	1990	29.3	41.7	44.3	213.8	125.5	90.7	171.8	145.1	195.5	143.5	36.5	29.9
Kenya	MWIRIM	MET	45	60.62	515	91340240	1991	6.8	27.4	159.9	183.8	205.9	62	156.7	118.1	101.5	126.2	92.6	72.9
Kenya	MWIRIM	MET	45	60.62	515	91340240	1992	26.1	150	100.5	261.9	138.1	58.8	84.9	145	52	59	74.8	35.4
Kenya	MWIRIM	MET	45	60.62	515	91340240	1993	68.6	12.9	130	128.6	152	129.1	187.1	195.1	46.6	156.6	49.8	8.4
Kenya	MWIRIM	MET	45	60.62	515	91340240	1994	0.6	10	45.6	184	72.3	127.9	154.3	140.7	164.9	258.4	32.6	78.5
Kenya	MWIRIM	MET	45	60.62	515	91340240	1995	52	152.2	14.2	75.4	271.8	111.1	86.5	116.8	116.2	61.9	77.5	21.9
Kenya	MWIRIM	MET	45	60.62	515	91340240	1996	0.4	13.5	97.4	160.6		130.3	180.9	208	27.8	72.9	185.7	29.2
Kenya	MWIRIM	MET	45	60.62	515	91340240	1997	13.5	59.7	119.6	133.1	130.7	112.7	167.6	92.9	138.3	149.6	96.6	34.8
Kenya	MWIRIM	MET	45	60.62	515	91340240	1998	180.4	66	140.3	223.7	105.8	68.2	79.2	201.9	310	47.9	63.7	5.8
Kenya	MWIRIM	MET	45	60.62	515	91340240	1999	8.5	0	160.3	257.2	84	116.3	159.1	121.3	8.5	152.2	105.3	74.7
Kenya	MWIRIM	MET	45	60.62	515	91340240	2000	163.5	59.4	17.5	170.1	140.9	84.6	176	148.7	82.8	137.5	97.8	8.3
Kenya	MWIRIM	MET	45	60.62	515	91340240	2001	6.8	3.3	86.2	181.3	95.5	151.2	118.5	99.4	74.5	275	66.3	21.9
Kenya	MWIRIM	MET	45	60.62	515	91340240	2002	19.3	0	13.6	126.9	154.1	74.1	145.2	184.6	88.2	154.9	56.4	12
Kenya	MWIRIM	MET	45	60.62	515	91340240	2003	53.4	10.7	169	113.6	163.9	184.2	168.3	95.7	108.9	194	169.5	8.1
Kenya	MWIRIM	MET	45	60.62	515	91340240	2004	9.7	32	129.4	158.3	115.6	99.8	78.9	146.2	36.4	150.3	45	146.7
Kenya	MWIRIM	MET	45	60.62	515	91340240	2005	22.4	6	87.7	285.8	224	189.9	165	188	72.9	26.1	46.7	58.2
Kenya	MWIRIM	MET	45	60.62	515	91340240	2006	27.4	28.8	73.8	190.4	123.9	60.6	76.5	142.2	56.9	147.7	135.5	60.8
Kenya	MWIRIM	MET	45	60.62	515	91340240	2007	44.6	17.9	42.7	139.2	178.6	100.6	81.2	152.5	131.8	146.7	30.1	11
Kenya	MWIRIM	MET	45	60.62	515	91340240	2008	13.2	66.5	119.3	247.1	108.3	60.9	79.7	133.2	221.7	160.5	202.7	78.8
Kenya	MWIRIM	MET	45	60.62	515	91340240	2009	46	110.4	74.6	221.6	167	158.3	186.2	187.2	286.5	64.1	29.1	37
Kenya	MWIRIM	MET	45	60.62	515	91340240	2010	6.9	26.3	60	99.8	154.1	128.7	122.9	249.8	96	167.6	103.1	0.7
Kenya	MWIRIM	MET	45	60.62	515	91340240	2011	36.7	6	24.1	259.1	141.8	35.1	99.4	118.4	89.8	135.9	74.5	233.2
Kenya	MWIRIM	MET	45	60.62	515	91340240	2012	45.3	137.4	170.3	216.3	213.4	91.2	197.3	88.9	113.3	163.1	49.5	20.3
Kenya	MWIRIM	MET	45	60.62	515	91340240	2013	2.8	19.7	167.6	138	176.7	88.3	108.2	265.5	141.1	76.3	301.5	34
Kenya	MWIRIM	MET	45	60.62	515	91340240	2014	0	7	45.2	260.9	189.9	110.6	193.5	74.7	258.3	157.4	159	71
Kenya	MWIRIM	MET	45	60.62	515	91340240	2015	41.4	119.9	122.7	146.7	310.7	161.8	156.6	264.9	62.6	82.9	22.3	153.4
							MEAN	31.7	45.5	95.3	182.9	172.1	108.4	132.8	155.1	130.3	121.1	94.0	46.5
							STDEV	40.8	46.2	67.8	58.4	59.3	44.6	43.1	48.5	121.3	64.0	68.0	48.1
							2STDEV	81.6	92.5	135.6	116.8	118.6	89.3	86.2	97.0	242.5	127.9	136.0	96.2
							LOWER LI	-49.9	-47.0	-40.4	66.1	53.5	19.2	46.6	58.0	-112.3	-6.8	-41.9	-49.8
							UPPER LIN	113.2	138.0	230.9	299.7	290.7	197.7	219.0	252.1	372.8	249.0	230.0	142.7
							20th PERC	6.7	9.4	40.7	128.3	124.9	68.1	84.2	114.1	55.9	60.8	36.4	11.8
						LOWER LI	**10th PERC**	2.89	4.11	23.75	105.62	106.30	61.23	79.35	96.81	44.36	46.71	30.46	8.16

Comparison with neighbors test

Cntr_Shift_ N; will sort data by year latitude longitude to compare neighbors.

Check if these high values are appearing in neighboring stations to confirm the consistence of the values identified and how they compare with neighbor in the similar years.

However this may not work in mountainous regions so other criteria checks can be employed such ENSO years or the temporal checks as par that very station time series data.

The Temperature quality control

The Temperature checks outside the database using the Macro and following the same steps highlighted above. The table below uses both the 2sgma upper limit and Step change from the previous day checks. The blue flags have been accepted but the one in red are still suspicious values set for further investigation

MINIMUM TEMPERATURE FLAGGED VALUES

Station	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	stdv	mean	Upperli	Lowerlimit
RAGA	150	150	150	160	130	150	150	160	170	190	170	150	150	110	110	23	154	199	108
WAU	210	200	200	210	220	200	210	210	220	220	230	210	190	190	170	17	204	238	170
JUBA	230	210	230	230	230	230	210	230	210	240	250	230	200	200	230	14	221	250	192
DJIBOUTI	200	180	202	197	199	205	201	196	190	215	230	234	210	200	215	16	212	243	181
LODWAR	211	214	215	225	219	216	215	215	231	222	232	215	216	220	228	19	234	271	196
MOYALE	214	208	211	215	200	215	212	212	196	202	160	195	196	220	212	10	206	225	186
MARSABIT	180	180	170	170	170	167	166	174	175	164	164	180	183	200	184	8	177	194	161
KITALE	114	96	115	115	116	129	130	125	114	114	118	111	105	125	125	10	121	141	101
WAJIR	225	225	226	232	235	235	236	232	240	230	225	230	235	225	225	5	232	242	221
ELDORET	106	145	135	120	100	123	140	112	109	91	79	96	88	104	118	18	111	148	75
KAKAMEGA	135	150	146	136	125	125	145	144	142	123	143	148	143	133	156	10	142	161	123
MERU	116	116	101	119	106	114	107	106	88	127	83	110	119	129	133	16	116	148	84
KISUMU	163	167	179	190	190	191	197	192	178	164	163	188	188	191	203	11	184	206	161
KISII	164	176	179	165	168	174	173	168	169	155	162	164	173	184	165	9	165	183	148
KERICHO	90	100	100	98	106	106	84	127	113	67	91	89	100	113	136	16	104	136	72
NAKURU	105	113	117	115	110	123	122	116	115	93	105	109	117	166	125	15	121	151	90
NYERI	83	84	79	87	90	116	91	106	75	115	91	99	115	110	119	18	104	139	68
EMBU	122	125	126	122	120	123	132	131	126	139	124	130	132	132	152	10	133	153	113
NAROK	999	999	89	95	86	110	82	91	98	74	80	95	94	128	154	235	169	640	-302
NRB/DAGORE	138	120	121	131	120	150	153	147	122	154	143	118	148	152	162	15	140	170	109
NRB/WILSON	143	137	145	138	142	168	165	135	142	158	155	138	167	147	155	13	153	180	127
MAKINDU	187	177	190	175	185	175	198	190	180	185	168	178	177	177	210	10	188	209	167
LAMU	244	237	240	243	227	243	246	254	226	219	226	244	214	252	256	12	240	264	215
VOI	210	207	215	211	211	214	207	210	211	205	195	208	203	204	224	9	214	233	195
MALINDI	256	253	249	254	257	240	256	257	240	238	231	243	242	264	245	9	248	266	231
MOMBASA	243	220	235	224	230	230	240	240	225	222	232	242	228	236	254	9	237	254	219
KAMEMBE	156	160	156	156	150	149	170	180	176	171	156	160	156	156	150	11	159	181	138
GISENYI	147	145	163	156	154	158	160	178	142	160	147	145	163	156	154	10	154	173	134
KIGALI	164	169	170	173	165	167	175	183	160	168	164	169	170	173	165	8	166	182	151
BUJUMBURA	193	200	184	181	192	198	224	224	207	187	187	195	200	204	214	13	193	219	167
Arua	141	137	151	146	160	166	151	154	157	146	150	151	167	144	142	9	148	167	129
GULU	201	221	220	210	205	200	200	235	212	212	222	210	215	196	181	11	208	229	186
Masindi	174	170	178	168	174	168	160	166	178	174	184	170	192	166	174	9	172	190	154
Kasese	155	148	167	160	165	170	180	159	180	170	170	162	176	170	161	10	167	187	148
Kampala	190	200	200	192	200	190	198	201	208	188	197	197	198	200	210	10	194	214	174
Jinja	165	174	177	179	174	164	154	163	163	168	203	168	170	163	199	13	172	197	147
Mbarara	172	165	181	171	170	170	170	170	180	169	174	176	172	176	179	8	169	185	152
Entebbe	195	204	206	209	205	195	199	207	219	199	198	202	200	199	224	13	197	224	171
Kabale	112	105	108	118	128	110	115	106	124	105	125	122	129	122	122	10	117	137	97
Kitgum	120	124	189	155	184	180	152	160	175	177	175	173	161	169	171	19	175	213	136

NMHS support

The NMHSs Data managers have been provided with these Macros and trained on error detection procedures. Tanzania Meteorological Authority (TMA) requested technical support to quality control their climate data sets (Rainfall and Temperature). A diagnostic approach to establish the error sources and causes is being done at ICPAC