Poway

Project Title: San Diego County Hydrograph Modification Plan   Project No: 133904

Date: February 2, 2009

Background

In consideration of other state and regional hydromodification management planning efforts, the County of San Diego has endeavored to compact at least 35 years of historical hourly rainfall data into several .wdm files toward the use of hydrologic modeling in the San Diego region. The hydrologic modeling software HSPF utilizes the .wdm files for both an input and output time series data (precipitation values over time). Each .wdm file is a binary, direct-access file that is organized into discrete data sets associated with either an ALERT or CCDA rainfall station. The ALERT stations are maintained by the County of San Diego; whereas, the CCDA stations are maintained by the Western Regional Climate Center and Scripps Institution of Oceanography's California Climate Change Center, thru funding by the California Energy Commission as part of the CalClim project. A total of eighteen ALERT stations and two CCDA stations were chosen to provide an adequate geographic coverage of the major watersheds within County.   This report will describe the methodology used toward the creation of the Poway ALERT station input .wdm file.

Alert Data – Framework and Data Gaps

The ALERT station at Poway (Sensor ID 24) is still in operation and has provided data between December 1962 and June 2008. The data set for any ALERT station is a series of line by line recordings which reflect the various wet or dry weather events logged by the gauge station. Each data set consists of the recorded sensor data as well as other "attributes" to describe the data (e.g. units).

During a wet event, the mechanical sampler within the gauge station would usually capture the amount of precipitation through the use of a tip bucket. Each time the bucket fills with rain, it tips see-saw like. The total rainfall is measured by counting how many times the bucket tips. However, this recording process was not uniformly done. The data within each ALERT station usually contains an assortment of logging periods at either a 5-minute, 15-minute, hourly or even daily interval. So while the mechanical sampling device is constantly recording the amount of precipitation, the line by line recordings within the time series data are exclusively subject to the whatever logging period was programmed. Additionally, these logging periods are used as the "Source" description in order to properly delineate each recorded event within the time series.

An example of the non-uniform data is provided in Table 1 below:

Table 1 – Poway County Data - RAW						
Date	Time	Data Value	TipBucketCount	Units	Code	Source
4/29/1983	730	10		in/100		Poway 15-min 1977-83
4/29/1983	745	10		in/100		Poway 15-min 1977-83
5/1/1983	615	10		in/100		Poway 15-min 1977-83
7/1/1983 1:06:00 AM		0	71	in	S1	Poway ALERT 1983-97
7/2/1983 1:06:00 PM		0	71	in	S	Poway ALERT 1983-97

However, a uniform dataset is necessary for use within the hydrologic modeling software. As such, the following standards were applied toward a uniform hourly dataset:

- If the ALERT data uses an hr 0, it is reflected in the final dataset as hr 24 of the previous day.

- If the ALERT data recorded rainfall precipitation in hundredths of an inch (in/100) or millimeters (mm), the precipitation value was corrected to inches and a "hundredths to inches" or "millimeters to inches" tagline is used within the source description.

- Aggregation – If an ALERT station recorded multiple rainfall entries within a distinct hour, such as when the logging program was set to 5-minute periods, the precipitation values of each entry are added together to reflect a final cumulative precipitation value for the hour. Regardless if multiple rainfall entries occurred or not, this aggregation process was used across all ALERT data entries due to such a large dataset. Once the aggregation process was applied, an "aggregated" tagline is inserted at the end of the source description.

- Disaggregation – If an ALERT station recorded a single daily rainfall entry, the data value provided is broken down into an hourly rainfall distribution. To accomplish this, a unit hyetograph was created following the direction of the County of San Diego Hydrology Manual, Section 4.3.2. The unit hyetograph uses a log-log interpolated value between the 5-year, 6-hr and 24-hr isopulvial graphs, which are also provided by the San Diego Hydrology Manual. Importantly, the disaggregation process generates a group of 24-hr precipitation distribution values which is often less than the initial cumulative daily data value provided due to rounding. If the disaggregation process was applied, a "disaggregated" tagline is inserted within the source description. The disaggreation table is included within Appendix A – Construction of Hyetograph.

Possible Inherent Errors – Data gaps

When creating the initial input .wdm, it is assumed that the recorded sensor data inherently contains errors or data gaps. After recording data for over three decades in a harsh outdoor environment, machinery can fail, sensors can be damaged or perhaps the electronic data was scrambled. Thankfully, the logging program which recorded the precipitation values would usually apply an error attribute alert during the two main data gap situations: an accumulation period or a missing period. An accumulation period is an error scenario where the amount of precipitation is known, but only as a single cumulative amount over a time frame beyond an hour. Basically the distribution of precipitation is unknown, rendering the data unproductive. (e.g. a faulty tip bucket). A missing period is an error scenario where the amount of precipitation is unknown over any time period. This occurred if something affected the logging apparatus directly (e.g. dead battery).

The identifier codes of an accumulation period are (but not limited to) the following: ***, C   SA, SA   A, A**

They are grouped together to display when the error period begins and ends. Occasionally, an error period would be followed by a subsequent error period resulting in back to back codes: **SA, SA, SA**

The identifier codes of a missing period are (but not limited to) the following: **M, -   SM, SM**

Possible Inherent Errors – Identify Outliers

Unfortunately, not all of these data gaps were so easily identified. Indeed, possible data gaps were at times given a 'passing' code, creating a scenario even more difficult to locate on record characteristics alone. As such, the outlier method of identifying errors was used. The general assumption for any time series data analysis is that it contains (1) a systematic pattern and (2) random event outliers (very high or very low data values) which will make the pattern difficult to identify. Unfortunately, a broad filter of all of these random events in order to make the pattern more salient cannot be implemented as many of these events were concretely observed. This is particularly true for all low data values, with the single exception of any possible negative precipitation events which would obviously be an error. However, some of the very high value outliers do indicate a data gap error. For example, an accumulation event will result in higher than normal precipitation values within a single hour.

BROWN AND CALDWELL

An example of the various errors discovered by the outlier method is shown in Table 2 below.

Code	DateMOD	HourMOD	Amount	Source	Units
			Table 2 - La Mesa County Data		
S	12/21/1989	17	0.08	La Mesa ALERT hrly 1983-97	in
*	12/21/1989	12	0	La Mesa ALERT hrly 1983-97	in
	12/26/1989	23	0.04	La Mesa ALERT hrly 1983-97	in
S	12/28/1989	7	0	La Mesa ALERT hrly 1983-97	in
SA	12/31/1989	24	0.12	La Mesa ALERT hrly 1983-97	in
S	1/31/1990	23	2.91	La Mesa ALERT hrly 1983-97	in
S	2/1/1990	15	0	La Mesa ALERT hrly 1983-97	in
S	2/4/1990	10	0.12	La Mesa ALERT hrly 1983-97	in
S	2/5/1990	10	0.54	La Mesa ALERT hrly 1983-97	in
S	2/6/1990	10	0.85	La Mesa ALERT hrly 1983-97	in
S	2/10/1990	10	0.22	La Mesa ALERT hrly 1983-97	in
S	2/25/1990	25	-0.02	La Mesa ALERT hrly 1983-97	in
SM	3/2/1990	4	0	La Mesa ALERT hrly 1983-97	in

Missing Code: C

Hidden Daily Data, not hourly

Negative event

Too large, SA event

Data Gap Analysis table

Trying to identify each data gap separately would be a time consuming process. Instead, the primary step toward the reestablishment of these data gap errors was to gather them together into a Data Gap Analysis table. As mentioned previously, data gaps occasionally occurred across a successive timeframe which became readily apparent once the gaps are grouped together. Conversely, a flagged data gap would occasionally occur within the same hour or subsequent hour. If the precipitation of the data gap is known (i.e. accumulation event), then the data gap warning can be ignored as the event either occurs within the same hour or subsequent hour. Unfortunately, such a straightforward scenario did not usually occur; as such, the time frame for each data gap is properly identified and then supplemented with verifiable data from a separate source. In order to provide a complete representation of this ALERT station, two historical online datasets were used to supplement any such data gaps. An appropriate supplemental gauge station from either of these historical datasets would be chosen on the basis of location proximity and similarity between the gauge and station elevations. However, a disparity between precipitation values recorded by the ALERT and NNDC publications will potentially exist no matter how precise the location/elevation similarity due to the disordered nature of a stormwater event. A listing of this numerical disparity and the selected supplemental gauge station per data gap is included within the Data Gap Analysis. The Data Gap Analysis table is included as Attachment B.

Filling in the Data Gaps - Supplemental Rainfall Record

NOAA

The first historical online dataset is freely available from the National Weather Service (NWS) at the following website: http://www.wrh.noaa.gov/sgx/obs/rtp/rtpmap.php?wfo=sgx. This link only provides recent daily precipitation data from Regional Temperature and Precipitation (RTP) stations within San Diego. Although not all of these stations reside in the same location as the ALERT stations, they cover some of the same general areas. In this case, the local RTP station at Station ID 648, Poway. If a data gap occurred within a recent timeframe (May 1998 or later), it would be supplemented by the daily data provided by NWS. Whenever a data gap is supplemented in such a fashion, a "NOAA" tagline is used within the source description as the NWS is operated by the National Oceanic Atmospheric Administration (NOAA).

BROWN AND CALDWELL

- ◆ Disaggregation – The daily data provided by *NOAA* is broken down into an hourly rainfall distribution to fit the time framework of the hydromodification plan. No matter the actual location of the NOAA station, the disaggregation table and unit hyetograph previously created for the Alert station is used. This unit hyetograph was created by following the direction of the County of San Diego Hydrology Manual, Section 4.3.2. It uses a log-log interpolated value between the 5-year, 6-hr and 24-hr isopulvial graphs, which are also provided by the San Diego Hydrology Manual. Importantly, the disaggregation process generates a total 24-hr precipitation distribution value which is less than the initial daily data value provided by NOAA due to rounding. If the disaggregation process was applied, a "disaggregated" tagline is inserted within the source description. The disaggreation table is included within Appendix A – Construction of Hyetograph.

NNDC

The second historical online dataset must be purchased from the National Climatic Data Center (NCDC), specifically thru the Image and Publications System at the following website: http://www7.ncdc.noaa.gov/IPS/hpd/hpd.html?_page=0&jsessionid=7F8ED28E1BD0C792DFBD4BA7 FAC08E51&state=CA&_target1=Next+%3E. This link provides a large volume of hourly precipitation data in PDF format from October 1951 through April 2008. A single PDF purchase yields the recorded hourly data of all NCDC stations within California for one month of a specific year. However, only database gaps up until May 1998 were supplemented in such a fashion. Any data gaps from May 1998 until the present were supplemented with the freely available daily data provided by NWS. Whenever a data gap is supplemented in such a fashion, a "NNDC" tagline is used within the source description to reflect the final data dissemination entity operated by the NOAA National Data Centers (NNDC).

Supplemental ALERT or CCDA stations

Whether thru design or coincidence, an occasional data gap would be found across all three historical datasets (ALERT, NOAA and NNDC). During such an unusual event, the supplemental information will be drawn from the various CCDA rainfall stations as each contains very minimal data gaps.

- ◆ The final source of supplemental information will be drawn from the nearest ALERT station. In this case, no additional ALERT stations were used.

QA/QC process

While the current goal is to gather together an input .wdm file, the final goal is to create an output .wdm time series analysis to forecast or predict future precipitation values. This predictive process depends upon identifying a sequence of observations (or pattern) within the observed time series data (i.e. input .wdm). Once the pattern is established, the data can be extrapolated to predict future events. As such, the focus of the QA/QC process for the input .wdm file is the filtering or removal of any errors which will interfere with the pattern, while forgiving any lesser errors which have little to no interference. Once the supplemental information was introduced, the primary QA/QC methods used were (1) a script to identify duplicate event times and (2) a repeat of the Outlier method.

Summary of Poway ALERT Station and .wdm Data Set

Watershed: **Los Penasquitos River**
Latitude/Longitude of Poway ALERT station: **32.5658, -117.0346**
Elevation of Poway ALERT Station: **440 feet**
Data available from: **10/04/1962 to 06/30/2008**
Highest Precipitation recorded within an hour: **0.80 inch**
Lowest Precipitation recorded within an hour: **0.00 inch**
Historical Dataset Resources: **Poway ALERT Data, NNDC Lake Wolford, NNDC Escondido No. 2, NOAA Poway, CCDA Lindbergh Field**

BROWN AND CALDWELL

APPENDIX A

Construction of Hyetograph

6-hr Storm

Hour #	Minutes (D)	Intensity	Total Rainfall	Rainfall by Hour	Poway	5-yr
1	60	0.30722816	0.30722816	0.30722816		
2	120	0.19642177	0.39284355	0.08561538		
3	180	0.15119873	0.45359619	0.06075264		
4	240	0.12557935	0.5023174	0.04872121		
5	300	0.10873686	0.54368432	0.04136693		
6	360	0.09666667	0.58	0.03631568		
	Sum			0.58		

6hr storm	24hr storm
1.75	3

Assumptions: The relationship between a 6hr and 24hr storm is 1:1.71 and that there are no rainfall depth-area adjustments.

**The San Diego County Manual, Section 4.3.2 was used to create the below rainfall distribution.**

For Hours 1-6; use I = 7.44 * (P6hr) * D^(-0.645)   and P=I * (D/60) where P6=0.58 inches and D is in minutes.
For Hours 6-24; use P(x)hr = (x^0.393) * (10^-0.54239)   where x equals a particular time in hours

m-slope =	0.393
b-intercept =	-0.54239

for hours 6-24 use the formula:
log [f(x)] = mlog(x) + b OR f(x)=x^m) * (10^b)

Nested Storm

Hour #	Minutes (D)	Intensity	Total Rainfall	Rainfall by Hour	Distribution
1	60	0.30722816	0.30722816	0.307	0.017
2	120	0.19642177	0.39284355	0.086	0.018
3	180	0.15119873	0.45359619	0.061	0.019
4	240	0.12557935	0.5023174	0.049	0.020
5	300	0.10873686	0.54368432	0.041	0.021
6	360	0.09666667	0.58	0.036	0.022
7	420		0.61622325	0.036	0.023
8	480		0.64942487	0.033	0.026
9	540		0.68019253	0.031	0.027
10	600		0.70894828	0.029	0.031
11	660		0.73600689	0.027	0.033
12	720		0.76161025	0.026	0.036
13	780		0.78594886	0.024	0.041
14	840		0.80917577	0.023	0.061
15	900		0.83141607	0.022	0.086
16	960		0.85277352	0.021	0.307
17	1020		0.87333523	0.021	0.049
18	1080		0.89317518	0.020	0.036
19	1140		0.91235681	0.019	0.029
20	1200		0.93093496	0.019	0.024
21	1260		0.94895744	0.018	0.021
22	1320		0.96646619	0.018	0.019
23	1380		0.98349825	0.017	0.018
24	1440		1.00008651	0.017	0.017
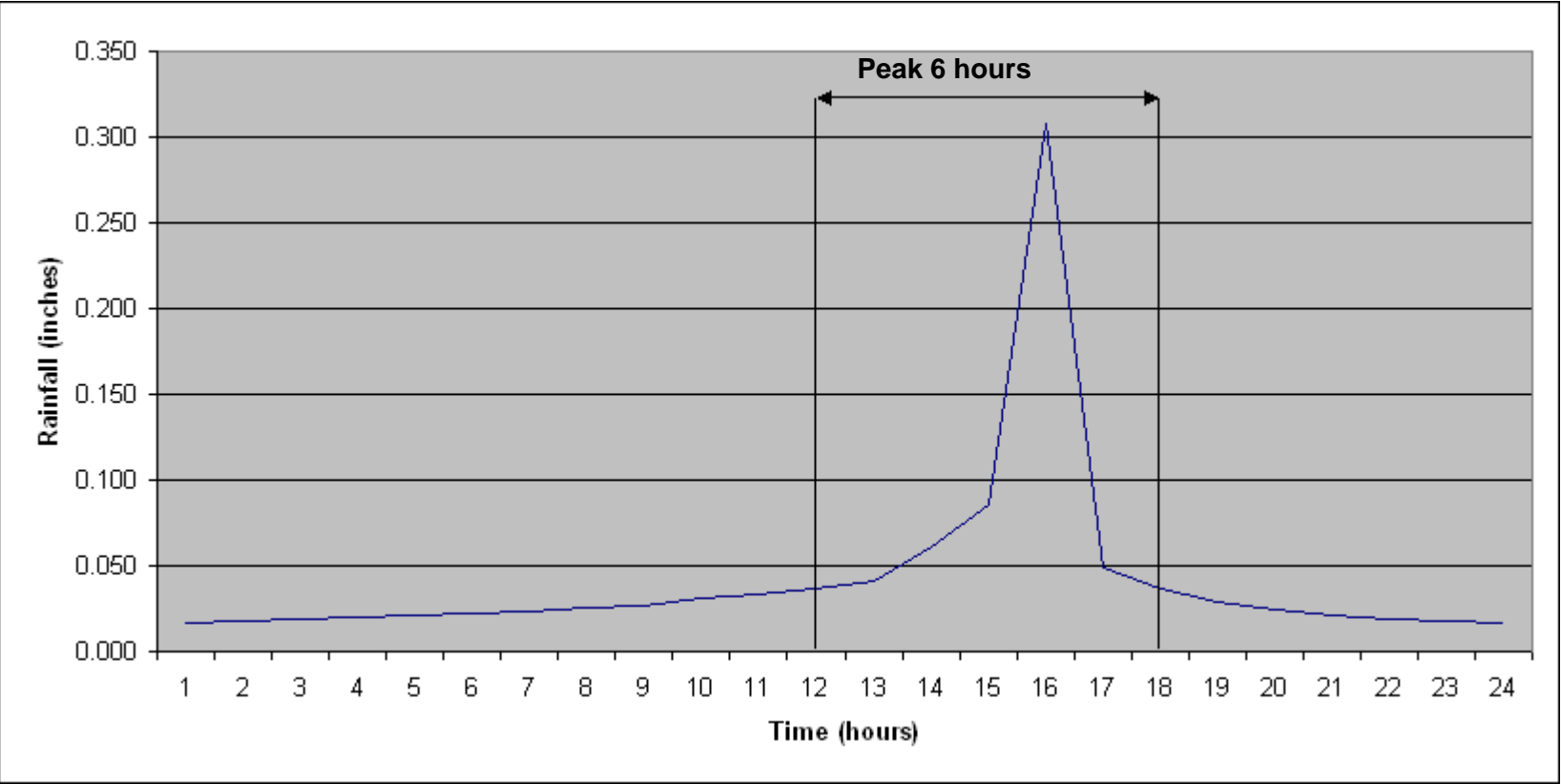
APPENDIX B

Data Gap Analysis Table

DateMOD	HourMOD	Amount	Code	Units	Filler Data	Value Comparison
7/1/1962	1	0	M	in	All nearby stations report no event	M
7/31/1962	24	0	-	in		
8/1/1962	1	0	M	in	All nearby stations report no event	M
8/31/1962	24	0	-	in		
9/1/1962	1	0	M	in	All nearby stations report no event	M
9/30/1962	24	0	-	in		
10/1/1962	1	0	M	in	NNDC Lake Wolford	M
10/31/1962	24	0	-	in		
2/9/1963	15	0	*	in	NNDC Lake Wolford	0.31
2/9/1963	24	1.22	C	in		
12/12/1971	1	0	*	in	NNDC Escondido No. 2	-0.05
12/15/1971	14	0.25	C	in		
1/16/1973	21	0	*	in	NNDC Escondido No. 2	-0.03
1/16/1973	24	0.63	C	in		
12/28/1976	9	0	*	in	NNDC Escondido No. 2	0.65
12/31/1976	24	0.75	C	in		
1/1/1977	1	0	*	in	NNDC Escondido No. 2	-0.46
1/4/1977	9	1.26	C	in		
4/5/1986	15	0	SA	in	NNDC Escondido No. 2	-0.16
4/6/1986	3	0.16	SA	in		
9/28/1987	23	0	SA	in	NNDC Lake Wolford	-1.26
9/30/1987	24	1.26	SA	in		
10/22/1987	19	0.04	SA	in	CCDA Lindbergh Field	-0.10
10/24/1987	20	0.16	SA	in		
11/4/1987	17	0	SA	in	CCDA Lindbergh Field	0.48
11/5/1987	12	0.43	SA	in		
4/26/1988	6	0	SA	in	NNDC Escondido No. 2	-0.04
6/30/1988	24	0.24	SA	in		
2/8/1989	1	0	SA	in	NNDC Escondido No. 2	-0.16
2/8/1989	13	0.16	SA	in		
12/31/1989	22	0	SA	in	CCDA Lindbergh Field	0.87
2/1/1990	12	1.65	SA	in		
6/8/1990	20	0	SA	in	CCDA Lindbergh Field	0.42
6/10/1990	4	0.35	SA	in		
3/13/1991	9	0	SA	in	CCDA Lindbergh Field	1.32
3/26/1991	24	3.78	SA	in		
2/12/1992	20	0.04	SA	in	CCDA Lindbergh Field	-0.01
2/13/1992	21	0.63	SA	in		
3/2/1992	7	0.04	SA	in	CCDA Lindbergh Field	0.19
3/2/1992	13	0.28	SA	in		
3/2/1992	14	0.04	SA	in	Ignore SA event, occurred in hour	0.00
3/2/1992	14	0.28	SA	in		
3/26/1992	18	0.04	SA	in	Ignore SA event, occurred in hour	0.00
3/31/1992	18	0	SA	in	CCDA Lindbergh Field	-0.29
3/31/1992	24	0.31	SA	in		
2/3/1994	20	0.08	SA	in		
2/3/1994	22	0.2	SA	in	CCDA Lindbergh Field	-0.24
2/4/1994	6	0.2	SA	in		
3/6/1994	24	0.35	SA	in	Ignore SA event, occurred in hour	0.00
12/5/1994	1	0	SA	in	Ignore SA event, occurred in hour	0.00

DateMOD	HourMOD	Amount	Code	Units	Filler Data	Value Comparison
12/24/1994	1	0	SA	in	CCDA Lindbergh Field	-0.14
12/25/1994	24	0.79	SA	in		
12/31/1994	13	0	SA	in	CCDA Lindbergh Field	-2.20
1/1/1995	1	2.2	SA	in		
1/3/1995	4	0	SA	in	CCDA Lindbergh Field	-0.03
1/3/1995	11	0.16	SA	in		
1/3/1995	13	0	SA	in	CCDA Lindbergh Field	-0.12
1/3/1995	24	0.2	SA	in		
1/4/1995	1	0	SA	in	CCDA Lindbergh Field	-0.40
1/4/1995	24	2.64	SA	in		
1/5/1995	1	0	SA	in	CCDA Lindbergh Field	-0.04
1/5/1995	24	0.16	SA	in		
1/7/1995	1	0	SA	in	CCDA Lindbergh Field	-0.80
1/7/1995	24	0.98	SA	in		
1/8/1995	1	0	SA	in	CCDA Lindbergh Field	-0.19
1/8/1995	24	0.35	SA	in		
1/11/1995	18	0	SA	in	CCDA Lindbergh Field	0.03
1/12/1995	8	0.2	SA	in		
1/25/1995	14	0	SA	in	CCDA Lindbergh Field	0.19
1/26/1995	19	0.83	SA	in		
12/31/1996	17	0	SA	in	CCDA Lindbergh Field	-1.83
1/15/1997	15	2.91	SA	in		
12/31/1999	24	0	SM	in	NOAA Poway and CCDA Lindbergh Field	M
3/9/2000	10	0	SM	in		
1/27/2001	3	0	SM	in	NOAA Poway	M
4/20/2001	10	0	SM	in		
8/8/2001	22	0	SM	in	NOAA Poway and CCDA Lindbergh Field	M
1/5/2002	10	0	SM	in		
12/29/2004	1	0.16	SA	in	Ignore SA event, occurred in hour	0.00
12/29/2004	1	0.04	SA	in		
1/9/2005	19	0.08	SA	in	Ignore SA event, occurred in 2nd hour	0.00
1/9/2005	20	0.39	SA	in		
2/26/2007	23	0	SA	in	NOAA Poway	-0.08
2/27/2007	20	0.24	SA	in		
2/28/2007	4	0	SA	in	NOAA Poway	0.20
2/28/2007	23	0.2	SA	in		
1/27/2008	2	0.12	SA	in	NOAA Poway	0.65
1/27/2008	6	0.16	SA	in		
1/27/2008	8	0.16	SA	in		