Is homogenisation of Australian temperature data any good?

Part 1. Methods case study: Parafield, South Australia

Dr Bill Johnston[1]
scientist@bomwatch.com.au

Abrupt up-steps in maximum temperature at Parafield airport near Adelaide S.A. in 1993 and 1999 were caused by site changes and there is no evidence that the climate has changed or warmed since 1940. Adjusting faulty Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) data using faulty Parafield data spreads problems across the ACORN-SAT network, which is used by the Bureau of Meteorology and CSIRO to estimate Australia's rate of warming.

Read on …

Summary

While Parafield airport (BoM ID 23013) is not an ACORN-SAT site, data was used by Bureau scientist Blair Trewin to adjust ACORN-SAT sites at Kent Town, Ceduna, North Shields (Port Lincoln), Nuriootpa, Oodnadatta, Robe, Snowtown, Tarcoola and Woomera in 2011. Woomera is used to adjust Maree, whose data are used to adjust Tibooburra, Alice Springs and Birdsville, Birdsville adjusts Charleville, Longreach and Richmond (Qld) …. Together with a wider range of cross-correlated second-tier stations, which for any one station likely embed parallel faults, the homogenisation process potentially spreads faults in Parafield data far and wide across the ACORN-SAT network.

In this Part 1 of a series examining ACORN-SAT homogenisation, Parafield is used as a case study to document methods developed by www.BomWatch.com.au to assess the fitness of maximum temperature (Tmax) datasets for determining trend and change in Australia's climate. The most striking difference between the BomWatch covariance approach and methods that examine data as time-series is that BomWatch deconstructs the temperature signal into component parts: one part due to site changes, which is investigated separately, and the other related to the climate. In contrast, time-series methods including naïve trend analysis, presumes the site has remained the same and that trend alone explains the data, which is not true.

As changes in ambience are detected using independent statistical methods procedures cannot be fiddled to achieve pre-determined outcomes. Multiple linear regression simultaneously verifies the effect of rainfall and site changes on Tmax and residuals are examined for unexplained systematic variation including climate-related trend. Bomwatch methods are physically based, objective, robust and replicable, which are cornerstones of the scientific method.

Adjusted for rainfall and step-changes, multiple linear regression residuals represent the 'true' climate signal and it is concluded that since records commenced in 1940, despite the increase of 2.0°C caused by site moves and changes, no trend or change in average Tmax at Parafield is attributable to the climate.

1. Introduction

Located about 15 km north of the city centre, Parafield airport was Adelaide's primary aerodrome until 1955. During WWII it served as a civilian airport and also as a training and transit base for the Royal Australian Air Force. The aerodrome serviced trans-Australia routes between the east coast and Perth,

---

[1] Former NSW Department of Natural Resources research scientist and weather observer.

and the main corridor from Adelaide to Darwin, and Aeradio, which vastly increased safety in the air, commenced operations there in 1939. While radio operators tracked aircraft, meteorological staff undertook regular observations including upper-atmosphere wind studies, made forecasts, provided pilot briefings and warned of inclement weather. Aeradio became a unit of the RAAF from 30 June 1940 until June 1946 when functions were split between Flight Services (radio communications, traffic control etc) and the Bureau of Meteorology (BoM).

Standard meteorological (met) observations usually included daily rainfall, rainfall intensity, daily maximum and minimum temperatures, terrestrial minimum temperature, regular (usually four-hourly) wet and dry-bulb thermometer readings for calculating relative humidity and dew point temperature, barometric pressure, visibility, surface and near surface (20 to 30 feet (10 m)) wind strength and direction, cloud type and octants (eights) of sky covered. A 63-degree cloud-base searchlight installed at Parafield in 1929 was removed in 1992 and evaporation was measured using an A-Pan evaporimeter from September 1975 to November 1989.

While wind strength, direction, visibility and cloud-height may have been observed from the roof of the control tower, other observations required an instrument enclosure be located near the runways within convenient distance of the met-office. As Aeradio commenced before the control tower was built, it must have operated from a separate building. An aerial view from 1948 shows a building of similar appearance to other Aeradio offices northeast of the tower, which is now used for some other purpose, and the probable location of the enclosure (Figure 1).

Figure 1. The imposing control tower and terminal at Parafield aerodrome 1948 and the probable location of the original Aeradio office (A) and meteorological enclosure (X). (National Archives of Australia.)

Coordinates for the enclosure were not provided in site-summary metadata. While the site may have moved a small distance to the north as the airport was redeveloped in the 1960s and 1970s, an aerial photograph shows that by November 1968 the Stevenson screen had relocated near the centre of the airport 1 km southwest of the control tower, 1.2 km north of the southern boundary, close to its present position (Figure 2). However, as few aerial photographs are available and metadata and other information is scant the time when the site moved to its current position is unknown. According to site summary metadata an automatic weather station (AWS) was installed at the current site in July 1992, which does not preclude that an AWS or analogue instrument operated previously somewhere else. Importantly, data held by the National Archives of Australia (NAA) from January 1955 to December 1972, which includes the change from Fahrenheit to Celsius on 1 September 1972 has apparently not been digitised. There is also no indication when the Parafield met-office closed; however, it could have been in 1989 when evaporation observations ceased. Typical of most Bureau sites, details of subsequent changes including upgrades and installation of a 60-litre screen have also not been noted in metadata.

Considering the scarcity of useful metadata, unreported station moves and changes which may or not have impacted on measurements, and the lack of corroborative information and photographs, analysis of trends and changes in maximum temperature data (Tmax) is best approached using objective statistical methods. This Report provides an overview of those methods using Parafield as a case study.

Figure 2. A rescaled overlay of a portion of the greyscale November 1968 South Australian Government aerial photograph and a Google Earth Pro satellite image for February 2019. Meteorological instruments are identifiable in both images immediately left of the point marked ‘x’. Note the airport was redesigned from its earlier ‘A’ configuration to parallel and offset runways and taxiways in the 1970s.

2. Methods

The First Law of thermodynamics, which states that energy is neither created nor destroyed, provides a physical reference frame against which the performance of individual weather stations can be assessed.

Briefly, energy available at the surface (R_S), which is the balance between incoming shortwave solar radiation minus reflection is dissipated via three pathways: (i), a portion is re-radiated as long wave emissions (R_L) in proportion to the 4th power of the absolute temperature (K) according to the Stefan-Boltzmann Law; (ii), a portion (LE) is used to do the work of evaporating water, where L is the latent heat of vaporisation (2.45MJ/kg) and E is the amount of water evaporated (kg); and, (iii), remaining sensible (measurable) heat is either advected to the local atmosphere (H) or stored in the landscape as potential energy (G) and dissipated later. According to the First Law Theorem the sum of the various transformations can be no more or less than the energy available at the surface:

$$R_S = R_L + LE + H + G \qquad \text{Equation (1)}$$

While energy balance fluxes occur at time-steps of milliseconds to minutes, the longer the time-step the more ‘averaged’ and interpretable are the various components. Seasonality resulting from the tilt of earth’s axis as it rotates around the sun is most apparent at monthly time-steps for example. At annual time-steps insolation (R_s) is relatively constant and as heating in summer is balanced by heat loss in winter the terms G and R_L effectively cancel out. Thus, for two days exactly 1-year apart net gains = losses and the balance is practically zero.

With components constant or accounted for, the local surface energy balance resolves into two terms: maximum temperature (Tmax), which measures heat advection from the landscape to the local atmosphere (H), and, evapotranspiration, the portion of R_s removed as latent heat via convection at the rate of 2.45MJ/kg of water evaporated. As 1 kg of water/m^2 = 1 mm rainfall, latent heat loss equates to 2.45MJ times rainfall equivalent. Thus, average rainfall at Parafield (447 mm) removes 447 * 2.45 = 1095 MJ/m^2 of heat from the local environment, or 17.23% of total solar exposure (6355 MJ/m^2). The deficit between incoming solar exposure and latent heat lost via evaporation, is dissipated as outgoing long-wave radiation and heat advection to the local atmosphere (Eqn (1)), which explains why the Mediterranean climate of South Australia cycles between cool moist winters and particularly warm dry summers.

As evaporation for dryland sites cannot exceed the rainfall, and for any weather station annual R_S is essentially fixed by Latitude, ignoring cancelled terms in Equation (1), it follows that a direct relationship is expected between heat advection (annual Tmax) and evaporation (annual rainfall), which can be determined statistically using linear regression as:

$$\text{Tmax} = a + b*\text{rainfall} + \epsilon \qquad \text{Equation (2)}$$

Where a is the intercept of the line with the Tmax-axis[1]; b the slope coefficient describing the rate of change of Tmax with rainfall (°C/100 mm) and ϵ is unexplained residual variation or statistical error. Consistent with the First Law Theorem, dry years are warm and the drier it is the warmer it gets.

While R_N depends on Latitude and annual rainfall depends on factors such as altitude, aspect, exposure to prevailing weather etc., the Tmax/rainfall coefficient is not constant. However, for a single site the coefficient is analogous to Bowen ratio (β=H/LE), which is used in micro and macro meteorology to partition sensible and latent heat fluxes in the context of estimating short-term (day, week, month) evaporation.

Using concurrent data and based on the First Law Theorem, methods developed by www.BomWatch.com.au uniquely pinpoint the effect of site changes on Tmax and verify their impact using straightforward statistical methods. The approach is applicable to weather stations spanning from the northern tropics to southern Tasmania. As protocols are physically based, rigorous, independent and cannot be fudged, lack of a relationship between Tmax and rainfall (as found for Cape Leeuwin and Portland Vic.) indicates problems with data or the site, but not the climate.

3. Detection of non-rainfall signals

Spreadsheet analysis of temperature time-series cannot discriminate between the signal of interest (Tmax) and underlying site-change effects (site inhomogeneities). Furthermore, practitioners of naïve methods rarely check that underlying assumptions are not violated (that residuals are homogeneous, independent, normally distributed with equal variance).

For example, annual average maximum temperature at Rutherglen Vic. shows an apparently significant linear trend of 0.61°C/decade ($P = 0.005$, $R^2_{adj} = 0.298$) (Figure 1). However, factors affecting the data including changed exposure of the site when it moved, introduction of an automatic weather station and changing to a small 60-litre screen caused the trend to be spuriously unrelated to the climate. Consequently residuals (data minus trend) are not random but exhibit systematic changes indicative of missing variable bias (bias due to a variable that has not been accounted for in the analysis). Zero-centred residuals show a suspicious decline from the start of the record to about 1960 and an increase from about 1980 to the end of the record, which indicates a problem in the data.

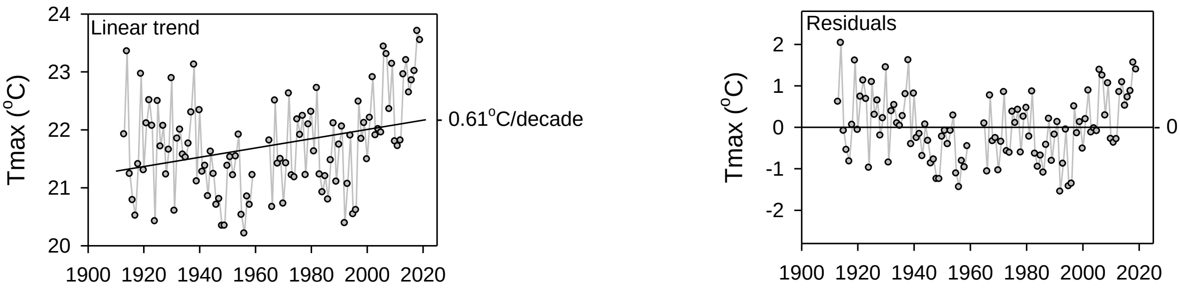

Figure 1. Tmax at Rutherglen appears to be increasing at a rate of 0.61°C/decade (left). However, due to site and instrument changes, zero-centred residuals (right) are not homogeneous, which invalidates the naïvely determined linear trend. Further analysis found the site had moved at least twice and that data were affected by the change to an automatic weather station in 1998, which was exacerbated by installation of a sensitive 60-litre Stevenson screen in 2006 (https://www.bomwatch.com.au/bureau-of-meteorology/rutherglen-victoria/).

As was the case for Rutherglen, maximum temperature time-series for Parafield consists of two overlaid parallel signals. Background variation caused by changes in the ambience of the site, which for

---

[1] When rainfall is zero, Tmax = $a + \epsilon$; so, the intercept (a) is also theoretically the highest possible value of Tmax.

unbiased data should remain constant, and relative to that, the annual average of day-to-day synoptic variation, which reflects the weather.

3.1 Decomposition using linear regression

Regression analysis partitions variation in the dependent variable (Tmax) caused by the physically related *deterministic* or explanatory variable (rainfall), from variation that is not explained by the relationship (i.e., residual or background variation). As rainfall removes heat from the environment, the *Tmax vs. rainfall* relationship is expected to be negative and statistically significant. The steepness of the coefficient (°C/100 mm of rainfall), its significance or P level and the proportion of variation explained (R^2_{adj}) assist in understanding the *goodness of fit*. Adjusted for the number of terms and cases, R^2_{adj} is a measure of quality of the dataset. A high-quality Tmax dataset would be highly significantly related to rainfall ($P << 0.05$), and R^2_{adj} would be high (>0.6 or 60% for instance).

In contrast, poor quality data from a poorly exposed or managed site would be indicated by a P level that was not different to zero ($P >> 0.05$), a near-zero coefficient and low explanatory power ($R^2_{adj} < 0.5$ or 50%). So, the first stage of the analysis aims to verify, that consistent with the First Law Theorem, Tmax is inversely related to rainfall, obtain baseline statistical parameters describing the Tmax ~ rainfall relationship and derive residuals from that relationship for further study.

While naïve linear regression removes the rainfall portion of the Tmax signal, second-stage analysis examines the non-rainfall residual portion (the ϵ in Equation (2)) for background changes that may be attributable to something else, particularly site related inhomogeneities that cause step-changes in the mean. To be clear, if rainfall residuals are independent, random and embed no further signals, variation in Tmax is fully explained and analysis is complete.

To assist in developing changepoint scenarios based on step-change analysis, inhomogeneities are cross-referenced to what is known about the site by examining BoM metadata, historical accounts in newspapers, aerial photographs etc. held for example by the National Archives and National Library of Australia, state archives and libraries, universities and collections held by museums and local history societies. As the least reliable source of information for BoM weather stations is their site-summary and ACORN-SAT metadata, investigation of some sites may not be easily resolved except by statistical methods.

3.2 Verification using covariate analysis

With segments specified as factors, covariate analysis (categorical multiple regression) simultaneously evaluates effects of site factors and rainfall on Tmax, and verifies that segmented responses to rainfall are the same (slopes are parallel) and that individual relationships are offset (that rainfall-adjusted segment means are different), which ensures that segmented regressions are not coincident. Alternative scenarios including ACORN-SAT changepoints can also be evaluated and results compared based on statistical outcomes and adjusted R^2 values or the Alkaike information criterion[1]. Importantly, for any dataset there is only one optimal outcome, which is that segmented regressions are statistically parallel and offset.

Having identified and attributed known sources of variation in Tmax data, a trend or change in the climate would be detectable as a significant, systematic residual signal that was unrelated to rainfall or site change effects.

In summary: regression analysis determines fitness of the dataset; residuals are analysed to detect changes in background ambience, while covariance analysis provides simultaneous verification and provides the basis of making unbiased adjustments to Tmax.

---

[1] https://en.wikipedia.org/wiki/Akaike_information_criterion

4. Case study Parafield

Tmax at Parafield appears to increase at a rate of 0.31°C/decade (Figure 3). The relationship also seems highly significant with *Year* explaining 62.9% of Tmax variation ($P < 0.001$, $R^2_{adj} = 0.692$). However, residuals are autocorrelated (not independent) and clustered into groups above and below the zero-centred line. Such patterns indicate an influential variable has not been accounted for, which invalidates the linear case. As two variables affect the data synchronously time-series analysis is unable to partition one from the other

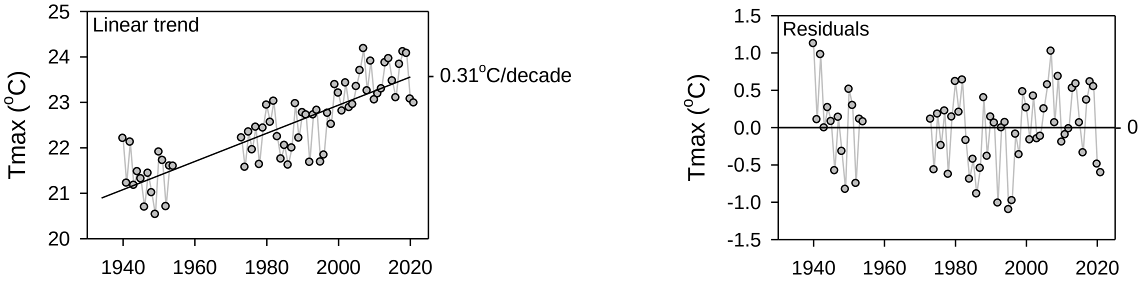

Figure 3. Tmax at Parafield appears to be increasing at a rate of 0.31°C/decade (left). However, although residuals are normally distributed, they are clustered above and below the zero-centred line (right). Residuals were therefore not independent ($P_{\text{no autocorrelation}} = 0.03$, Durbin-Watson test) nor random ($P_{\text{random}} = 0.01$, runs test), which invalidates naïvely determined trend.

Tmax at Parafield is significantly negatively correlated with rainfall (-0.398°C/100 mm, $P < 0.001$) and rainfall naïvely explains 16% of Tmax variation (Figure 3(a)). Residuals from that relationship (Tmax minus the rainfall effect, re-scaled by the grand mean) tested positive to up-steps in 1973, when after a gap of 17 years observations re-commenced, and in 1999 when it is likely a more sensitive 60-litre Stevenson screen replaced the former 230-litre one at the same site (Figure 3(b)). Breaking the data into raw data segments (Figure 3(c) to (e)) provides further insights including that data from the vicinity of the control tower to 1954 were probably affected by watering (the coefficient is not significant) and that averages step-up from 22.3°C in 1973 and 23.4°C (rainfall adjusted) in 1999.

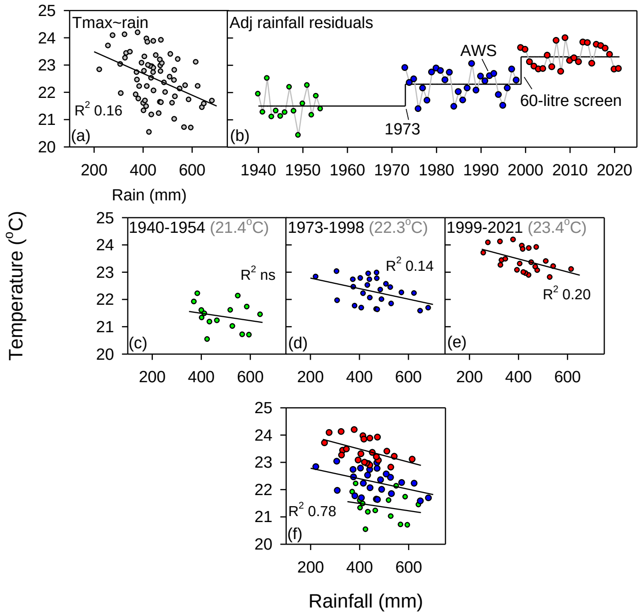

Figure 3. Composite analysis of Parafield Tmax (see text).

Categorical multiple linear regression (Figure 3 (f)) confirmed that: (i), rainfall-adjusted group means were offset (lines in (c) to (e)) were not coincident); (ii), responses to rainfall were the same across groups (interaction was not significant thus slopes were the same); (iii), that rainfall reduced Tmax at the rate of 0.2°C/100 mm ($P < 0.001$) and (iv), despite misgivings in relation to early data, step-changes and rainfall simultaneously explained 78% of Tmax variation.

Furthermore, rainfall domain residuals were independent, normally distributed with equal variance across groups. (Note: Data used in the study are provided in the appendix.)

Data files in the National Archives show that although data are unavailable from the BoM and there is no indication where the site was, observations continued during the data gap from 1955 to 1972. Data from the BoM recommenced from December 1972, the site moved before 1968 (Figure 2), and some equipment was decommissioned on 22 July 1992 on the same day an AWS was installed. Metadata holds the key to understanding the data but like most other sites, site-summary metadata supplied by BoM is incomplete and unreliable.

Supplementary analysis found no trend, partial trends or other irregularities that could be attributed to factors aside from rainfall and site-related changes. It follows that multiple linear regression residuals represent the 'true' climate signal and it is concluded that despite the apparent increase of 2.0°C since records commenced, no trend or change is attributable to the climate.

4. Conclusions

On both sides of the global warming divide, analysis of time series without accounting for site changes invariably leads to false perceptions, claims and counter-claims that are indefensible and therefore not robust. Analysis methods developed and used by www.BomWatch.com.au provide five levels of certainty in the outcome:

- Consistent with the First Law Theorem, a significant negative relationship is expected between average annual Tmax and rainfall. Step-change analysis is a critical test for determining homogeneity of residuals and the effect of underlying step-changes on trend.
- Step-changes scenarios are evaluated and compared successively using categorical multiple linear regression (MLR) with the aim of identifying the scenario that best explains the data (i.e., the scenario that minimises the residual sum of squares and the Alaike information criterion and maximises R^2_{adj}). Scenarios are also cross-referenced to available metadata and other sources including aerodrome maps and plans and aerial photographs. However, as sites may change without affecting data, and metadata may be incomplete or faulty, independent statistical tests should hold precedence over other information sources.
- MLR verifies that rainfall-adjusted segment means are different (segmented regressions are not coincident) and that responses to rainfall are the same (interaction is not significant, thus slopes are parallel), which validates the step-change test.
- Finally, segmented data are analysed independently in rainfall and time domains for trend and irregularities including outliers and bad data.
- Regression residuals are expected to be independent, normally distributed with equal variance and not affected by inhomogeneities.

Having accounted for site changes and rainfall, residual trend due to other factors such as CO_2 should and would be detectable in residuals. Also, as analysis is objective and statistical, it cannot be fudged or manipulated to create trends that do not exist. Finally, the scheme outlined here can be used to evaluate the veracity of HQ and ACORN-SAT changepoint scenarios applied to the same data.

Dr Bill Johnston

25 November 2022

Preferred citation:

Johnston, Bill 2022. Is homogenisation of Australian temperature data any good? Part 1. Methods case study: Parafield Airport South Australia. http://www.bomwatch.com.au/ 9 pp.

Appendix. Data used in the study

Row	Year	Rain	MaxAve	MaxN	MaxVar	ShiftFactor ShSite	ShiftFactor ShMaxRes	ShiftFactor ShBadMax
1	1940	386.2	22.20	365	51.52	1	1	0
2	1941	465.7	21.21	362	41.34	1	1	0
3	1942	551.2	22.12	365	43.86	1	1	0
4	1943	435.6	21.17	364	49.00	1	1	0
5	1944	414.4	21.47	365	45.92	1	1	0
6	1945	403.8	21.31	365	37.90	1	1	0
7	1946	597	20.69	362	38.14	1	1	0
8	1947	642.2	21.43	364	42.29	1	1	0
9	1948	529	21.01	363	40.05	1	1	0
10	1949	426.4	20.53	364	33.71	1	1	1
11	1950	371.7	21.90	365	38.97	1	1	0
12	1951	588.1	21.72	365	52.79	1	1	0
13	1952	569	20.70	366	37.84	1	1	0
14	1953	520.3	21.59	365	44.91	1	1	0
15	1954	402.8	21.59	365	39.55	1	1	0
16	1955			0		1	1	0
17	1956			0		1	1	0
18	1957			0		1	1	0
19	1958			0		1	1	0
20	1959			0		1	1	0
21	1960			0		1	1	0
22	1961			0		1	1	0
23	1962			0		1	1	0
24	1963			0		1	1	0
25	1964			0		1	1	0
26	1965			0		1	1	0
27	1966			0		1	1	0
28	1967			0		1	1	0
29	1968			0		1	1	0
30	1969			0		1	1	0
31	1970			0		1	1	0
32	1971			0		1	1	0
33	1973	624.6	22.21	365	47.29	2	2	0
34	1974	649.8	21.56	365	41.07	2	2	0
35	1975	487.6	22.34	365	44.82	2	2	0
36	1976	311	21.95	366	41.18	2	2	0
37	1977	377	22.44	365	47.99	2	2	0
38	1978	470.4	21.63	364	46.29	2	2	0
39	1979	528.2	22.43	365	48.65	2	2	0
40	1980	438	22.93	365	45.25	2	2	0
41	1981	510.6	22.55	365	52.40	2	2	0
42	1982	308.6	23.02	365	56.18	2	2	0
43	1983	573.8	22.24	365	50.35	2	2	0
44	1984	382.8	21.75	366	42.51	2	2	0
45	1985	444.4	22.04	365	43.18	2	2	0
46	1986	475.4	21.61	365	46.18	2	2	0
47	1987	492	21.99	365	44.77	2	2	0
48	1988	472.2	22.96	366	42.82	2	2	0
49	1989	417.4	22.21	365	54.23	2	2	0
50	1990	406	22.77	365	46.52	2	2	0
51	1991	375	22.72	365	43.84	2	2	0
52	1992	682.7	21.68	356	39.98	2	2	0

53	1993	443	22.71	351	45.84	3	2	0
54	1994	223.2	22.82	354	49.61	3	2	0
55	1995	409.6	21.68	364	50.04	3	2	1
56	1996	532	21.83	363	45.37	3	2	1
57	1997	473.4	22.75	357	53.91	3	2	0
58	1998	434.6	22.51	355	50.11	3	2	0
59	1999	513.6	23.38	344	54.83	3	3	0
60	2000	543.5	23.20	349	57.63	3	3	0
61	2001	529.4	22.80	362	57.27	3	3	0
62	2002	332.8	23.42	360	46.08	3	3	0
63	2003	443.8	22.88	363	53.19	3	3	0
64	2004	431	22.94	365	50.04	3	3	0
65	2005	453.8	23.34	363	41.71	3	3	0
66	2006	258.8	23.69	365	50.79	3	3	0
67	2007	380	24.18	364	53.92	3	3	0
68	2008	328.6	23.25	365	50.15	3	3	0
69	2009	474.6	23.90	365	60.76	3	3	0
70	2010	478.4	23.05	363	49.72	3	3	0
71	2011	470.4	23.18	363	40.66	3	3	0
72	2012	407.2	23.29	366	50.89	3	3	0
73	2013	444.2	23.86	365	48.42	3	3	0
74	2014	416	23.95	365	50.80	3	3	0
75	2015	348.4	23.46	364	58.12	3	3	0
76	2016	616.6	23.09	365	47.83	3	3	0
77	2017	419.6	23.83	364	49.07	3	3	0
78	2018	326.6	24.11	364	56.01	3	3	0
79	2019	277.4	24.07	364	64.50	3	3	0
80	2020	396.2	23.06	364	45.14	3	3	0
81	2021	422	22.98	365	41.01	3	3	0