

CONDITIONAL CLIMATOLOGY GIVEN THE LOCAL OBSERVATION

Robert G. Miller

**Techniques Development Laboratory
Systems Development Office
National Weather Service, NOAA
Silver Spring, MD 20910**

GEM

CONDITIONAL CLIMATOLOGY GIVEN THE LOCAL OBSERVATION

Robert G. Miller

Techniques Development Laboratory
Systems Development Office
National Weather Service, NOAA
Silver Spring, MD 20910

Editor's Note: Some of the principles that I have followed in selecting articles to be published in the Digest are that they be practical and written so that they appeal to the majority of the readership of the Digest. I have long been familiar with Dr. Miller's work in the Techniques Development Laboratory, NWS, and have often asked Bob to write an article for the Digest because, I felt, his work would greatly add to the advancement of practical meteorology. Dr. Miller's GEMTRIX is a novel, powerful idea, yet easy to understand and can be adapted to a forecasting scheme, particularly when applied to AFOS. The Digest thanks Dr. Miller for his article and invites your comments on it.

Randy Racer
Managing Editor

Brief Biographical Sketch

The author is Chief of the Objective Forecast Branch in the National Weather Service's Techniques Development Laboratory. He has spent almost thirty years developing and applying statistical methods to weather forecasting. His career began as a World War II weather observer serving in India and Burma. Prior to receiving his M.S. in meteorology from New York University he was an observer on ocean station vessels. His research work began at Massachusetts Institute of Technology and continued at the Travelers Research Center, Bell Telephone Laboratories, and the Life Insurance Marketing and Research Association. Prior to joining TDL he was Chief Scientist of the Air Weather Service. He holds a Ph.D. from Harvard in statistics and is the author of numerous technical papers including an AMS Meteorological Monograph entitled "Statistical Prediction by Discriminant Analysis".

INTRODUCTION

Classical climatology gives an average picture of the weather. It can be made more and more specific by conditioning on the basis of time and

location. The ultimate conditional climatology would be the one that specifies the probability distribution of all elements everywhere, conditioned on the full state of the atmosphere both past and present. A more practical conditional climatology would be one that takes into account the set of all locally observed meteorological variables to arrive at future conditional probability distributions. This paper will demonstrate the application of such a conditional climatology called GEM, an acronym for Generalized Equivalent Markov model.

METHODOLOGY

The approach will be to estimate the conditional probability distribution of every weather element one hour into the future, given the present condition of those same weather elements. Estimation of the probabilities is performed by multiple linear regression equations. To accomplish this, historical observations are used in the regression analysis. These are schematically represented in Table 1.

The X's and Y's shown in Table 1 represent, for example, month of the year, hour of the day, temperature, ceiling height, visibility, weather, cloud cover, and so forth.

Next, the observations of the X's and Y's are transformed into intervals (or categories). Each is assigned a 1.0 or 0.0 depending, respectively, upon whether that observation is in or out of the interval (or category). This transformation process serves two purposes. First, the regression equations will then actually estimate the desired probabilities (see Miller 1964). Second, the method to extend the conditional probability-distribution estimates beyond one hour is greatly facilitated. Consequently, each column in Table 1 expands into two or more columns which are mutually exclusive and exhaustive. A complete list of the transformations in this paper is given in Table 2.

Table 1. Schematic representation of an historical data sample of N observations, for the initial weather-variable conditions $X_{n,p}$ and the subsequent weather-variable conditions $Y_{n,p}$ where the subscript n denotes the general observation and the subscript p denotes the general variable.

Observation	Initial Weather Variables			Subsequent Weather Variables				
	X_1	X_2	X_p	Y_1	Y_2		Y_p	
1	$x_{1,1}$	$x_{1,2}$...	$x_{1,p}$	$y_{1,1}$	$y_{1,2}$...	$y_{1,p}$
2	$x_{2,1}$	$x_{2,2}$...	$x_{2,p}$	$y_{2,1}$	$y_{2,2}$...	$y_{2,p}$
.								
.								
.								
n	$x_{n,1}$	$x_{n,2}$...	$x_{n,p}$	$y_{n,1}$	$y_{n,2}$...	$y_{n,p}$
.								
.								
.								
N	$x_{N,1}$	$x_{N,2}$...	$x_{N,p}$	$y_{N,1}$	$y_{N,2}$...	$y_{N,p}$

A 10-year historical sample of data, consisting of 86,782 hourly observations for Washington National Airport (DCA), was organized in the fashion described above. Five years (the odd numbered years) were held aside for subsequent independent tests. The 5 even numbered years became the sample from which the regression equations were developed. The resulting 175 regression equations constitute a matrix of 175 columns and rows. For demonstration purposes, the set of all cloud amount equations is given in Table 3, where the 175 predictors are those identified in Table 2.

To appreciate the effect any variable has on, say, the overcast event, look at the coefficient of that variable. Since the variables are all either 1.0 or 0.0, the coefficient amount is either added (algebraically) or not added. Thus, if the observation is for January (variable number 2), then .02676120 is added to the conditional probability that overcast will be the observed condition one hour hence, all other things held constant. Since one of the other months must be a 1.0 if January is not "on", the situation where it is really August (variable number 9) causes -.01619480 to be added algebraically. This is a swing of .04295600 in the conditional probability of overcast.

Given an observation of all the weather elements at DCA at a particular time, the set of 175 regression equations is then used to estimate the 175 conditional probabilities one hour hence. To arrive at estimates of the 175 conditional probabilities two hours hence the original observation is replaced by the conditional probabilities estimated for one hour. This process of replacing the observation with the previous estimate may be continued ad infinitum. Obviously, the estimates of the conditional probabilities will reflect the fact that only seasonal and diurnal effects are of importance out beyond a certain period of time. The initial observation will eventually lose all of its influence. The process of iteration is referred to as the equivalent Markov process (see Miller 1968 and Whiton 1977).

Another important requirement of a useful conditional climatology is that it be convertible into "best" categorical estimates. This suggests applying a thresholding method that yields certain desired effects, such as producing categorical frequencies that are close to observed frequencies. A method developed by Miller and Best (1979) has been employed here, where the necessary parameters are the readily available multiple correlation coefficients and sample climatologies.

Table 2. Each X and Y in Table 1 has been transformed into the intervals and categories shown in this Table. Each element has been numbered for future reference. The number one element is always unity.

2	MONTH (LOCAL)	JAN	51	13 to 14
3		FEB	52	15 to 17
4		MAR	53	18 to 22
5		APR	54	23 to 27
6		MAY	55	28 to 32
7		JUN	56	33 to 45
8		JUL	57	SLP (MB) 980.1 to 990.0
9		AUG	58	990.1 to 1000.0
10		SEP	59	1000.1 to 1005.0
11		OCT	60	1005.1 to 1010.0
12		NOV	61	1010.1 to 1020.0
13		DEC	62	1020.1 to 1030.0
14	HOUR (LOCAL)	00	63	1030.1 to 1090.0
15		01	64	DB TEMP (DEG F) -99 to 14
16		02	65	15 to 23
17		03	66	24 to 27
18		04	67	28 to 29
19		05	68	30 to 31
20		06	69	32
21		07	70	33 to 34
22		08	71	35 to 36
23		09	72	37 to 38
24		10	73	39 to 40
25		11	74	41 to 42
26		12	75	43 to 50
27		13	76	51 to 59
28		14	77	60 to 68
29		15	78	69 to 77
30		16	79	78 to 86
31		17	80	87 to 140
32		18	81	DPT DEPRESSION 0
33		19	82	(DEG F) 1
34		20	83	2 to 4
35		21	84	5 to 6
36		22	85	7 to 11
37		23	86	12 to 15
38	WIND DIRECTION	CALM	87	16 to 99
39		NNE to NE	88	SKY COVER CLR
40		ENE to E	89	SCD
41		ESE to SE	90	BKN
42		SSE to S	91	OVC
43		SSW to SW	92	VISIBILITY .00 to .49
44		WSW to W	93	(MI) .50 to .74
45		WNW to NW	94	.75 to .99
46		NNW to N	95	1.0 to 1.49
47	WIND SPEED	0 to 3	96	1.5 to 1.99
48	(KT)	4 to 7	97	2.0 to 2.49
49		8 to 9	98	2.5 to 2.99
50		10 to 12	99	3.0 to 3.99

100	VISIBILITY (CONT)	4.0 to 4.99	154	-19 to -10
101		5.0 to 5.99	155	-9 to 0
102		6.0 to 6.99	156	1 to 9
103		7.0 to 99.00	157	10 to 19
104	WEATHER	NO WX	158	20 to 29
105		WX	159	30 to 39
106		NO F,IF	160	40 to 900
107		F,IF	161	CIG (100's FT) 0 to 1
108		NO GF	162	2 to 4
109		GF	163	5 to 6
110		NO K,H,D,KH,KD,HD,KHD	164	7 to 9
111		K,H,D,KH,KD,HD,KHD	165	10 to 14
112		NO BS,BD,BN,BY	166	15 to 19
113		BS,BD,BN,BY	167	20 to 24
114		NO L	168	25 to 29
115		L-	169	30 to 39
116		NO R	170	40 to 49
117		R-	171	50 to 59
118		R	172	60 to 74
119		R+	173	75 to 99
120		NO RW	174	100 to 150
121		RW-	175	151 to UNL
122		RW		
123		RW+		
124		NO S,IC		
125		S-,IC-		
126		S,IC		
127		S+,IC+		
128		NO SW,IP,SP		
129		SW-,IP-,SP-		
130		SW,IP,SP		
131		NO ZL		
132		ZL-		
133		NO ZR		
134		ZR-		
135		NO T,A-		
136		T,A-		
137	CLOUD COVER	CLR <200 FT		
138		OVC <200 FT		
139		CLR <500 FT		
140		SCD <500 FT		
141		BKN <500 FT		
142		OVC <500 FT		
143		CLR <1000 FT		
144		SCD <1000 FT		
145		BKN <1000 FT		
146		OVC <1000 FT		
147		CLR <3000 FT		
148		SCD <3000 FT		
149		BKN <3000 FT		
150		OVC <3000 FT		
151	PRES CHG (TENTHS	-900 to -40		
152	OF MB)	-39 to -30		
153		-29 to -20		

Table 3. Regression equations for the four cloud amount categories. The number to the left of each regression coefficient corresponds to the predictor number provided in Table 2.

88. SKY CVR CLR	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
1.095	1.229	1.347	1.459	1.574	1.693	1.816	1.943	2.074	2.209	2.348	2.491	2.638	2.789	2.944	3.103	3.266	3.433	3.604	3.779	3.958	4.141	4.328	4.519	4.714	4.913	5.116	5.323	5.534	5.749	5.968	6.191	6.418	6.649	6.884	7.123	7.366	7.613	7.864	8.119	8.378	8.641	8.908	9.179	9.454	9.733	10.016	10.303	10.594	10.889	11.188	11.491	11.798	12.109	12.424	12.743	13.066	13.393	13.724	14.059	14.398	14.741	15.088	15.439	15.794	16.153	16.516	16.883	17.254	17.629	18.008	18.391	18.778	19.169	19.564	19.963	20.366	20.773	21.184	21.599	22.018	22.441	22.868	23.3	23.731	24.157	24.588	25.023	25.462	25.905	26.352	26.803	27.258	27.717	28.18	28.647	29.118	29.593	30.072	30.555	31.042	31.533	32.028	32.527	33.03	33.537	34.048	34.563	35.082	35.605	36.132	36.663	37.198	37.737	38.28	38.827	39.378	39.933	40.492	41.055	41.622	42.193	42.768	43.347	43.93	44.517	45.108	45.703	46.302	46.905	47.512	48.123	48.738	49.357	49.98	50.607	51.239	51.875	52.516	53.161	53.81	54.464	55.121	55.782	56.447	57.116	57.789	58.466	59.147	59.832	60.521	61.214	61.911	62.612	63.317	64.026	64.739	65.456	66.177	66.902	67.631	68.364	69.101	69.842	70.587	71.336	72.089	72.846	73.607	74.372	75.141	75.914	76.691	77.472	78.257	79.046	79.839	80.636	81.437	82.242	83.051	83.864	84.681	85.502	86.327	87.156	87.989	88.826	89.667	90.512	91.361	92.214	93.071	93.932	94.797	95.666	96.539	97.416	98.297	99.182	100.071	100.964	101.861	102.762	103.667	104.576	105.489	106.406	107.327	108.252	109.181	110.114	111.051	111.992	112.937	113.886	114.839	115.796	116.757	117.722	118.691	119.664	120.641	121.622	122.607	123.596	124.589	125.586	126.587	127.592	128.601	129.614	130.631	131.652	132.677	133.706	134.739	135.776	136.817	137.862	138.911	139.964	141.021	142.082	143.147	144.216	145.289	146.366	147.447	148.532	149.621	150.714	151.811	152.912	154.017	155.126	156.239	157.356	158.477	159.602	160.731	161.864	163.001	164.142	165.287	166.436	167.589	168.746	169.907	171.072	172.241	173.414	174.591	175.772	176.957	178.146	179.339	180.536	181.737	182.942	184.151	185.364	186.581	187.802	189.027	190.256	191.489	192.726	193.967	195.212	196.461	197.714	198.971	200.232	201.497	202.766	204.039	205.316	206.597	207.882	209.171	210.464	211.761	213.062	214.367	215.676	216.989	218.306	219.627	220.952	222.281	223.614	224.951	226.292	227.637	228.986	230.339	231.696	233.057	234.422	235.791	237.164	238.541	239.922	241.307	242.696	244.089	245.486	246.887	248.292	249.701	251.114	252.531	253.952	255.377	256.806	258.239	259.676	261.117	262.562	264.011	265.464	266.921	268.382	269.847	271.316	272.789	274.266	275.747	277.232	278.721	280.214	281.711	283.212	284.717	286.226	287.739	289.256	290.777	292.302	293.831	295.364	296.901	298.442	300.000	301.557	303.118	304.683	306.252	307.825	309.402	310.983	312.568	314.157	315.75	317.347	318.948	320.553	322.162	323.775	325.392	327.013	328.638	330.267	331.9	333.533	335.17	336.814	338.461	340.112	341.767	343.426	345.089	346.756	348.427	350.102	351.781	353.464	355.151	356.842	358.537	360.236	361.939	363.646	365.357	367.072	368.791	370.514	372.241	373.972	375.707	377.446	379.189	380.936	382.687	384.442	386.201	387.964	389.731	391.502	393.277	395.056	396.839	398.626	400.417	402.212	404.011	405.814	407.621	409.432	411.247	413.066	414.889	416.716	418.547	420.382	422.221	424.064	425.911	427.762	429.617	431.476	433.339	435.206	437.077	438.952	440.831	442.714	444.601	446.492	448.387	450.286	452.189	454.096	456.007	457.922	459.841	461.764	463.691	465.622	467.557	469.496	471.439	473.386	475.337	477.292	479.251	481.214	483.181	485.152	487.127	489.106	491.089	493.076	495.067	497.062	499.061	501.064	503.071	505.082	507.097	509.116	511.139	513.166	515.197	517.232	519.271	521.314	523.361	525.412	527.467	529.526	531.589	533.656	535.727	537.802	539.881	541.964	544.051	546.142	548.237	550.336	552.439	554.546	556.657	558.772	560.891	563.014	565.141	567.272	569.407	571.546	573.689	575.836	577.987	580.142	582.301	584.464	586.631	588.802	590.977	593.156	595.339	597.526	599.717	601.912	604.111	606.314	608.521	610.732	612.947	615.166	617.389	619.616	621.847	624.082	626.321	628.564	630.811	633.062	635.317	637.576	639.839	642.106	644.377	646.652	648.931	651.214	653.501	655.792	658.087	660.386	662.689	664.996	667.307	669.622	671.941	674.264	676.591	678.922	681.257	683.596	685.939	688.286	690.637	692.992	695.351	697.714	700.081	702.452	704.827	707.206	709.589	711.976	714.367	716.762	719.161	721.564	723.971	726.382	728.797	731.216	733.639	736.066	738.497	740.932	743.371	745.814	748.261	750.712	753.167	755.626	758.089	760.556	763.027	765.502	767.981	770.464	772.951	775.442	777.937	780.436	782.939	785.446	787.957	790.472	792.991	795.514	798.041	800.572	803.107	805.646	808.189	810.736	813.287	815.842	818.401	820.964	823.531	826.102	828.677	831.256	833.839	836.426	839.017	841.612	844.211	846.814	849.421	852.032	854.647	857.266	859.889	862.516	865.147	867.782	870.421	873.064	875.711	878.362	881.017	883.676	886.339	889.006	891.681	894.36	897.043	899.73	902.424	905.119	907.818	910.521	913.228	915.939	918.654	921.373	924.096	926.823	929.554	932.289	935.028	937.771	940.518	943.269	946.024	948.783	951.546	954.313	957.084	959.859	962.638	965.421	968.208	970.999	973.794	976.593	979.396	982.203	985.014	987.829	990.648	993.471	996.298	999.129	1001.964	1004.803	1007.646	1010.493	1013.344	1016.199	1019.058	1021.921	1024.788	1027.659	1030.534	1033.413	1036.296	1039.183	1042.074	1044.969	1047.868	1050.771	1053.678	1056.589	1059.504	1062.423	1065.346	1068.273	1071.204	1074.139	1077.078	1080.021	1082.968	1085.919	1088.874	1091.833	1094.796	1097.763	1100.734	1103.709	1106.688	1109.671	1112.658	1115.649	1118.644	1121.643	1124.646	1127.653	1130.664	1133.679	1136.698	1139.721	1142.748	1145.779	1148.814	1151.853	1154.896	1157.943	1160.994	1164.049	1167.108	1170.171	1173.238	1176.309	1179.384	1182.463	1185.546	1188.633	1191.724	1194.819	1197.918	1201.021	1204.128	1207.239	1210.354	1213.473	1216.596	1219.723	1222.854	1225.989	1229.128	1232.271	1235.418	1238.569	1241.724	1244.883	1248.046	1251.213	1254.384	1257.559	1260.738	1263.921	1267.108	1270.299	1273.494	1276.693	1279.896	1283.103	1286.314	1289.529	1292.748	1295.971	1299.198	1302.429	1305.664	1308.903	1312.146	1315.393	1318.644	1321.899	1325.158	1328.421	1331.688	1334.959	1338.234	1341.513	1344.796	1348.083	1351.374	1354.669	1357.968	1361.271	1364.578	1367.889	1371.204	1374.523	1377.846	1381.173	1384.504	1387.839	1391.178	1394.521	1397.868	1401.219	1404.574	1407.933	1411.296	1414.663	1418.034	1421.409	1424.788	1428.171	1431.558	1434.949	1438.344	1441.743	1445.146	1448.553	1451.964	1455.379	1458.798	1462.221	1465.648	1469.079	1472.514	1475.953	1479.396	1482.843	1486.294	1489.749	1493.208	1496.671	1500.138	1503.609	1507.084	1510.563	1514.046	1517.533	1521.024	1524.519	1528.018	1531.521	1535.028	1538.539	1542.054	1545.573	1549.096	1552.623	1556.154	1559.689	1563.228	1566.771	1570.318	1573.869	1577.424	1580.983	1584.546	1588.113	1591.684	1595.259	1598.838	1602.421	1606.008	1609.599	1613.194	1616.793	1620.396	1623.903	1627.414	1630.929	1634.448	1637.971	1641.498	1645.029	1648.564	1652.103	1655.646	1659.193	1662.744	1666.299	1669.858	1673.421	1676.988	1680.559	1684.134	1687.713	1691.296	1694.883	1698.474	1702.069	1705.668	1709.271	1712.878	1716.489	1720.104	1723.723	1727.346	1730.973	1734.604	1738.239	1741.878	1745.521	1749.168	1752.819	1756.474	1760.133	1763.796	1767.463	1771.134	1774.809	1778.488	1782.171	1785.858	1789.549	1793.244	1796.943	1800.646	1804.353	1808.064	1811.779	1815.498	1819.221	1822.948	1826.679	1830.414	1834.153	1837.896	1841.643	1845.394	1849.149	1852.908	1856.671	1860.438	1864.209	1867.984	1871.763	1875.546	1879.333	1883.124	1886.919	1890.718	1894.521	1898.328	1902.139	1905.954	1909.773	1913.596	1917.423	19

National Weather Digest
VERIFICATION RESULTS

Applying the methodology described above to the 5 years of independent data produced the verification statistics shown in Table 4 for 1-h conditional-probability estimates.* Additional verification tests have yielded favorable results for GEM at six other locations around the country, for both 1- and 6-h projections (see also Miller, Whiton, and Kelly 1977).

Table 4. Comparative verification for six ceiling and visibility categories between classical climatology, conditional persistence, and the GEM conditional climatology model. Shown are the Brier scores for 43,047 1-h independent projections (lower values are better).

Brier Scores		
	Ceiling	Visibility
Classical Climatology	.43709	.36571
Conditional Persistence	.20579	.19006
GEM conditional climatology	.19249	.17240

where the six categories are defined as:

Ceiling	Visibility
1. < 200'	1. < 1/2 mile
2. 200' ≤ to < 500'	2. 1/2 mile ≤ to < 1 mile
3. 500' ≤ to < 1000'	3. 1 mile ≤ to < 3 miles
4. 1000' ≤ to < 3000'	4. 3 miles ≤ to < 5 miles
5. 3000' ≤ to < 7500'	5. 5 miles ≤ to < 7 miles
6. ≥ 7500'	6. ≥ 7 miles

AN EXAMPLE

The derivation of a particular conditional climatology given the local observation will normally be performed automatically in a minicomputer such as an AFOS Eclipse. For demonstration purposes, an application of the methodology described above will now be made for a particular DCA observation. The steps performed are given below and in the flowchart shown in Figure 1.

- Step 1:** Collect input observation and transform into 1.0's or 0.0's depending upon observed interval or category.
- Step 2:** Estimate the 1-h conditional probabilities for each element using regression equations. Iterate out hour by hour to twenty-four hours using equivalent Markov principle, thereby creating a GEMTRIX (A Generalized Equivalent Markov Probability Matrix).
- Step 3:** Convert the probabilities in GEMTRIX into "best" hourly categorical estimates

*Classical climatology is defined as the dependent sample relative frequencies, and conditional persistence is represented by the dependent sample relative frequencies conditioned on the currently observed ceiling or visibility category.

using threshold probabilities, thereby creating a CATRIX (A GEMTRIX converted to a Categorical matrix).

- Step 4:** Assume hourly categorical estimates to be what will actually be observed and create an FT either subjectively or objectively.
- Step 5:** Forecaster then accepts or modifies CATRIX through feedback after studying FT and additional current weather information. A final FT is ultimately issued.

PERFORMING STEP 1:

The DCA observation chosen for this example is for 1600 local standard time (2100 GMT) on June 17, 1979 and was as follows:

Month : June

Hour : 1600 LST

Wind Direction : 330°

Wind Speed : 10 Kt

Sea Level Pressure : 1011.2 mb

Temperature : 74°F

Dew Point Depression (T-T_d) : 1°F

Sky Cover : Overcast

Visibility : 7 mi

Weather : None

Cloud Cover : Overcast under 3000'

Three Hour Pressure Change : -.7 mb

Ceiling : 1100 ft

The corresponding transformed observed values are shown in the first column of the GEMTRIX of Table 5.

PERFORMING STEP 2:

The estimated conditional probability distributions for all weather elements, projected from one to 24 hours, are given in Table 5 in the form of a GEMTRIX. The probabilities have been truncated and renormalized when necessary to have them lie between 0.0 and 1.0 and their sums equal to 1.0. In Table 5 the actual observed interval or category is parenthesized, when not obvious, for purposes of visual verification.

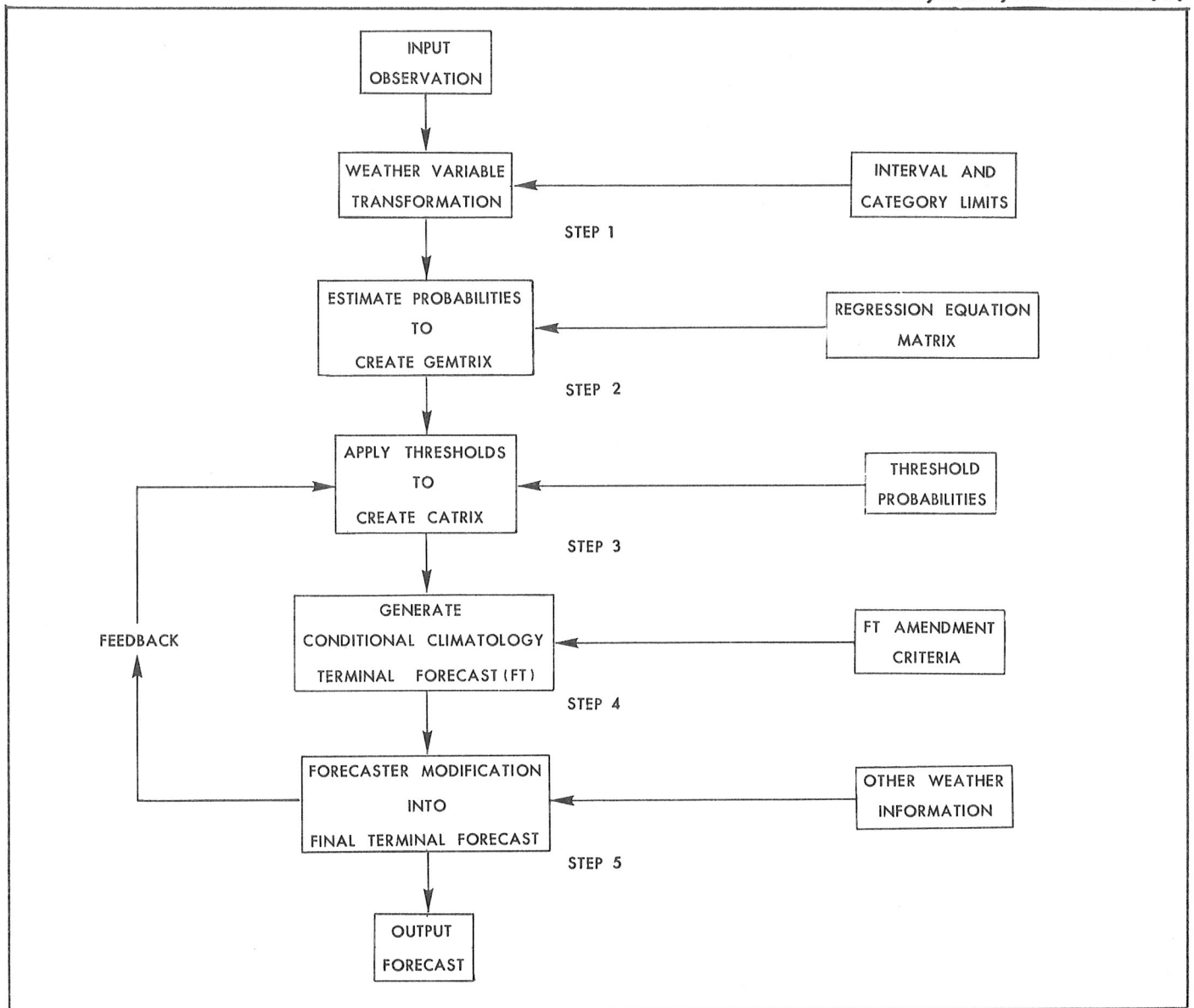


Figure 1. Flowchart of steps to follow in applying GEM to making an aviation terminal forecast.

PERFORMING STEP 3:

Table 6 gives the CATRIX for each weather element, again from one to 24 hours. The numbers in the CATRIX represent the "best" category. These are identified by using Table 2, starting with one if the first category, two if the second category, etc. for each weather element. Thus, a 4 for SKY COVER at hour 23 LST means the fourth category of SKY COVER or OVC.

PERFORMING STEP 4:

Assume the categories shown in the CATRIX will actually occur. A minicomputer algorithm could be written to produce an FT using amendment criteria. However, for this example two experienced forecasters generated the climatological FT shown below from only the information contained in the CATRIX of Table 6.

FORECASTER'S CLIMATOLOGICAL FT

DCA FT 8 SCT C12 OVC 3410.

05Z 10 SCT C15 OVC 6 H.

11Z C15 OVC 3FH.

17Z C15 BKN 5H BKN V SCT 1010.

21Z MVFR CIG BCMG VFR.

NNNN


```

XXXX XXX XXXXX XXXX X X
X XXXX X X X X X
X XXXX X X X X X
XXXX X X X XXXX X X
    
```

CATEGORICAL GEMTRIX
TECHNIQUES DEVELOPEMENT LABORATORY

FOR STATION 'DCA'

VALID FROM 79061717 TO 79061816
BASED ON 79061716 SINGLE STATION SURFACE OBSERVATION

CATEGORY	OBS	16	17	18	19	20	21	22	23	24	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
MONTH (LOCAL)	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	
MONTH DIR	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
WIND SPD (KTS)	9	4	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
SUP (MB)	2	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
DPT TEMP (DEG F)	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
SKY CVR (ST MI)	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
WEAT IFR	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
FOG IFR	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
GROUND FOG	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
KS.H.D.KH.KD.HD.KHD	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
KS.H.D.KH.KD.HD.KHD	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
DRN	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
RAIN	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
SNOW	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
SHOWERS	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
ICW	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
ICWERS	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
DRIZZLE	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
FREEZING	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
STORM	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
THUNDER	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
CLD < 200 FT	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
CLD < 500 FT	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
CLD < 1000 FT	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
CLD < 3000 FT	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
CPRES	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
CHG (.1 FT)	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
CIG (100'S FT)	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11

Table 6. Categorical matrix (CATRIX) generated from GEMTRIX in Table 5 from threshold probabilities designed to create distributions that agree with the observed distributions. The numbers in each cell are identified by referring to the intervals or categories in Table 2.

Table 7. Proposed weather element transformation intervals and categories to succeed those in Table 2.

1	MONTH (LOCAL)	JAN	51	10-19KTS
2		FEB	52	NNW-N 1 - 9KTS
3		MAR	53	10-19KTS
4		APR	54	NE \geq 20KTS
5		MAY	55	SE \geq 20KTS
6		JUN	56	SW \geq 20KTS
7		JUL	57	NW \geq 20KTS
8		AUG	58	SLP (MB) SLP \leq 985.0
9		SEP	59	985.1 to 990.0
10		OCT	60	990.1 to 995.0
11		NOV	61	995.1 to 1000.0
12		DEC	62	1000.1 to 1005.0
13	HOUR (LOCAL)	00	63	1005.1 to 1010.0
14		01	64	1010.1 to 1015.0
15		02	65	1015.1 to 1020.0
16		03	66	1020.1 to 1025.0
17		04	67	1025.1 to 1030.0
18		05	68	1030.1 to 1035.0
19		06	69	1035.1 to 1040.0
20		07	70	1040.1 to 1090.0
21		08	71	DB TEMP -130 to -31
22		09	72	(DEG F) -30 to -26
23		10	73	-25 to -21
24		11	74	-20 to -16
25		12	75	-15 to -11
26		13	76	-10 to -6
27		14	77	-5 to -1
28		15	78	0 to 4
29		16	79	5 to 9
30		17	80	10 to 14
31		18	81	15 to 19
32		19	82	20 to 24
33		20	83	25 to 29
34		21	84	30 to 34
35		22	85	35 to 39
36		23	86	40 to 44
37	WIND ROSE	CALM	87	45 to 49
38		NNE-NE 1 - 9KTS	88	50 to 54
39		10-19KTS	89	55 to 59
40		ENE-E 1 - 9KTS	90	60 to 64
41		10-19KTS	91	65 to 69
42		ESE-SE 1 - 9KTS	92	70 to 74
43		10-19KTS	93	75 to 79
44		SSE-S 1 - 9KTS	94	80 to 84
45		10-19KTS	95	85 to 89
46		SSW-SW 1 - 9KTS	96	90 to 94
47		10-19KTS	97	95 to 99
48		WSW-W 1 - 9KTS	98	100 to 104
49		10-19KTS	99	105 to 109
50		WNW-NW 1 - 9 KTS	100	110 to 114

101	DB TEMP	115 to 150 ^o F	157	HGT (100's FT)	0 to 1
102	DPT DEPRESSION	0	158		2 to 4
103	(DEG F)	1	159		5 to 6
104		2 to 4	160		7 to 9
105		5 to 7	161		10 to 14
106		8 to 11	162		15 to 19
107		12 to 15	163		20 to 24
108		16 to 19	164		25 to 29
109		20 to 25	165		30 to 39
110		26 to 35	166		40 to 49
111		36 to 50	167		50 to 59
112		51 to 115	168		60 to 75
113	SKY1 PARTIAL OBSCURATION		169		76 to 99
114	TOTAL OBSCURATION		170		100 to 150
115	CLR	0/10	171		≥ 151
116	SCD	1/10 to 5/10	172	VISIBILITY	0 to .49
117	BKN	6/10 to 9/10	173	(MI)	.50 to .74
118	OVC	10/10	174		.75 to .99
119	HGT1 (100's FT)	0 to 1	175		1.00 to 1.49
120		2 to 4	176		1.50 to 1.99
121		5 to 6	177		2.00 to 2.49
122		7 to 9	178		2.50 to 2.99
123		10 to 14	179		3.00 to 3.99
124		15 to 19	180		4.00 to 4.99
125		20 to 24	181		5.00 to 5.99
126		25 to 29	182		6.00 to 6.99
127		30 to 39	183		7.00 to 100.00
128		40 to 49	184	WEATHER	NO WX
129		50 to 59	185		F,IF
130		60 to 75	186		GF
131		76 to 99	187		K,H,D,KH,KD,HD,KHD
132		100 to 150	188		BS,BD,BN,BY
133		≥ 151	189		L-
134	SKY2 CLR	0/10	190		L
135	SCD	1/10 to 5/10	191		L+
136	BKN	6/10 to 9/10	192		R-
137	OVC	10/10	193		R
138	HGT2 (100's FT)	0 to 1	194		R+
139		2 to 4	195		RW-
140		5 to 6	196		RW
141		7 to 9	197		RW+
142		10 to 14	198		S-,IC-
143		15 to 19	199		S,IC
144		20 to 24	200		S+,IC+
145		25 to 29	201		SW-,IP-,SP-
146		30 to 39	202		SW,IP,SP
147		40 to 49	203		SW+,IP+,SP+
148		50 to 59	204		ZL-
149		60 to 75	205		ZL
150		76 to 99	206		ZL+
151		100 to 150	207		ZR-
152		≥ 151	208		ZR
153	SKY3 CLR	0/10	209		ZR+
154	SCD	1/10 to 5/10	210		T LIGHT
155	BKN	6/10 to 9/10	211		T HEAVY
156	OVC	10/10	212	PRES CHNG	-90.0 to -4.0
				(MB)	

Table 7.
continued

213	PRES CHNG (CONT)	-3.9 to -3.0
214		-2.9 to -2.0
215		-1.9 to -1.0
216		-.9 to 0.0
217		.1 to .9
218		1.0 to 1.9
219		2.0 to 2.9
220		3.0 to 3.9
221		4.0 to 90.0

PERFORMING STEP 5:

The same forecasters were then provided the most recent synoptic charts. As a result they made one modification to the climatological FT they originally created. Namely, they added "Chance of Thundershowers" for 05Z. Thus the final FT became:

FORECASTER'S FINAL FT

DCA 8 SCT C12 OVC 3410.

05Z 10 SCT C16 OVC CHC TRW

11Z C15 OVC 3FH.

17Z C15 BKN 5H BKN V SCT 1010.

21Z MVFR CIG BCMG VFR.

NNNN

COMMENTS ON EXAMPLE

Favorable:

- o Weather can actually be initiated and dissipated when employing the CATRIX. This, in fact, occurred in the example at numerous times, e.g., fog, haze, and their dissipation.
- o Note that it was possible to enter "probabilities" for the month to split the effect of the observation not being at midmonth. The same thing could have been done if the time was a fractional part of the hour.
- o In general there is a high degree of consistency among the weather elements in CATRIX throughout the 24-h period, e.g., between visibility, dewpoint depression, and the types of weather anticipated.

Unfavorable:

- o Snow probabilities greater than 0.0 in June with temperatures in the 60's is disconcerting. This is due to there being no guarantee that adding terms in the regression equation, for certain input conditions, will yield acceptable probability estimates (technically referred to as the nonadditivity problem). The saving grace in this situation is that it did not, and would not likely, appear in CATRIX since CATRIX acts as a filter. This problem will be completely eliminated with the new statistical procedure called Discrete Likelihood Functions (DLF) because the probability under these initial conditions would come out to be 0.0 (see Miller 1979).
- o Lack of consistency between the overall no weather variable and its component parts. However, this is correctable. In this example, the proper hierarchy would have been to ignore the NO WX variable and only allow for weather in the CATRIX when a particular type of weather, such as fog or haze, is categorically estimated to occur.
- o Ceiling conditions proved to be estimated on the pessimistic side relative to the actual outcome although a certain amount of fluctuation was observed.
- o The intervals and categories chosen here do not appear to be optimal. Our future plans call for revamping them.

SUMMARY AND FUTURE WORK

It would seem that the procedure described and demonstrated here would be useful as an aid to the forecaster especially at short range, say, one

to six hours. In particular, it has the following beneficial characteristics:

1. More skillful than climatology and persistence where independent verifications have shown GEM to be as good as or better than TAP (Terminal Alert Procedure; see Vercelli and Hefferman 1978) for ceiling and visibility. Thus, GEM would seem to be a logical successor to TAP for AFOS since it includes projections on all of the remaining weather elements.
2. Capable of giving interesting and useful "best" category estimates as demonstrated in the "creation" and "dissipation" of fog and haze.
3. Since only the current observation is required to use GEM, any significant changes encountered in the present weather can initiate an instantaneous update of future conditions. There is no need to wait for the record observation. It is anticipated that this will be extremely well suited to the AFOS era both as a monitoring device and as a highly refined climatological aid to the forecaster.
4. Needs only a minicomputer such as exists in AFOS to perform the necessary calculations.
5. Is fully objective with interpretable regression coefficients as can be seen by studying the cloud cover equations in Table 3.
6. Provides a "what if" capability for the forecaster by merely modifying the input.

Future plans call for the following refinements:

1. Defining better intervals and categories in the weather variable transformations. For example, see Table 7.
2. Developing a generalized operator system so that the same regression equations are applicable anywhere provided an observation and that location's climatology are available. This is made possible by extracting location climatology at the outset, before developing the generalized regression equations, and then having it reinstated to make the equation "location specific".
3. Mapping forecasts when these are available for the entire area. This would give a capability for analysis and "time-lapse" projection.
4. Accounting for interactions among the observed variables by means of a technique such as DLF will be tested.
5. Performing a spectral decomposition of the regression-equation matrix employing eigen-

value analysis. This should yield fundamental climatological concepts and interpretations.

6. Obtaining best single estimates in CATRIX for elements such as temperature in degrees F or C. These can and will be gotten by using expectation or some similar procedure.

ACKNOWLEDGEMENTS

This work has been performed with the help of Albert Forst, William Griner, William Herrmann, and Frank Lewis. The author wishes to sincerely thank each of these individuals for his unusual perseverance and skillful performance. Many enjoyable and fruitful hours were spent in conceptual discussions with Larry Best to whom the writer is grateful. Thanks are also given to Mary B. Battle for typing the manuscript.

REFERENCES

- Miller, R. G., 1964: Regression estimation of event probabilities. Tech. Report 7411-121, Contract Cwb 10704, Travelers Research Center, Inc., Hartford, Conn., 153 pp.
- _____, 1968: A stochastic model for real-time on-demand weather predictions. Proceedings of the First Statistical Meteorological Conference, Hartford, Amer. Meteor. Soc., 48-51.
- _____, 1979: Estimating event probabilities by discrete likelihood functions. Preprints Sixth Conference on Probability and Statistics in Atmospheric Sciences, Banff, Amer. Meteor. Soc., 93-97.
- _____, and D. L. Best, 1979: A model for converting probability forecasts to categorical forecasts. Preprints Sixth Conference on Probability and Statistics in Atmospheric Sciences, Banff, Amer. Meteor. Soc., 98-102.
- _____, R. C. Whiton, and M. J. Kelly, Jr., 1977: Results of a single station forecasting experiment. Preprints Fifth Conference on Probability and Statistics in Atmospheric Sciences, Las Vegas, Amer. Meteor. Soc., 37-40.
- Vercelli, D. J., and M. M. Hefferman, 1978: AFOS aviation local monitoring and updating program. Preprints Conference on Weather Forecasting and Analysis and Aviation Meteorology, Silver Spring, Amer. Meteor. Soc., 369-373.
- Whiton, R. C., 1977: Markov processes in Selected Topics in Statistical Meteorology, edited by R. G. Miller, Air Weather Service Tech. Report AWS-TR-77-273, Chapter 7, 45 pp.