

《中国与大国关系数据库》使用方法

本数据库采用定量衡量方法，将 1950 年以来的中国与美国、日本、俄罗斯（苏联）、英国、法国、印度和德国这七个大国的双边关系用分值表示出来，现将本书使用方法介绍如下。

双边关系分值的等级

在图 1—1 “双边关系分值标准”中，双边关系划分为三大类别：“敌对”、“非敌非友”、“友善”，每一类别按照程度再一分为二，共六个等级，对应的分值范围分别是：“对抗”（-9—-6）、“紧张”（-6—-3）、“不和”（-3—0）、“普通”（0—3）、“良好”（3—6）、“友好”（6—9）。每个等级再分为三等水平：低等水平、中等水平和高等水平，正值与负值的低、中、高方向相反。例如，在良好等级中，低等水平是 3—4，中等水平是 4—5，高等水平是 5—6；在紧张等级中，低等水平是 -3—-4，中等水平是 -4—-5，高等水平是 -5—-6。每个水平分为十度，每一度的分值为 0.1。在本书中，0.1 是最小衡量单位。例如，处于良好等级高等水平的双边关系可能处于 5.1、5.2、5.3……5.9 的任何一点上。

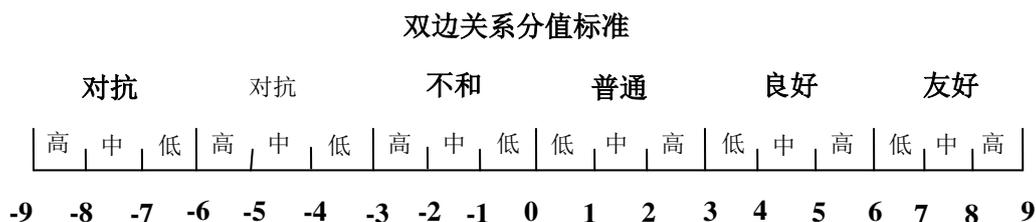


图 1—1

如何使用图形

一、单一双边关系比较

单一双边关系比较是通过观察同一双边关系曲线在不同时期的位置、方向、振频和变化幅度所进行的比较。观察曲线位置可比较关系的好坏程度，观察曲线方向可比较关系的走势，观察曲线的振频可比较关系的稳定性，观察同长时间内曲线的变化幅度可比较关系的变化速度。下文将以中美关系的变化曲线为例做上述四方面的观察和比较。

1、比较同一双边关系在不同时期的友善或敌对程度。例如，观察图 1—2 中的中美关系曲线，读者可知中美关系在 1979 年和 1988 年都处于普通等级，但 1988 年的关系好于 1979 年。

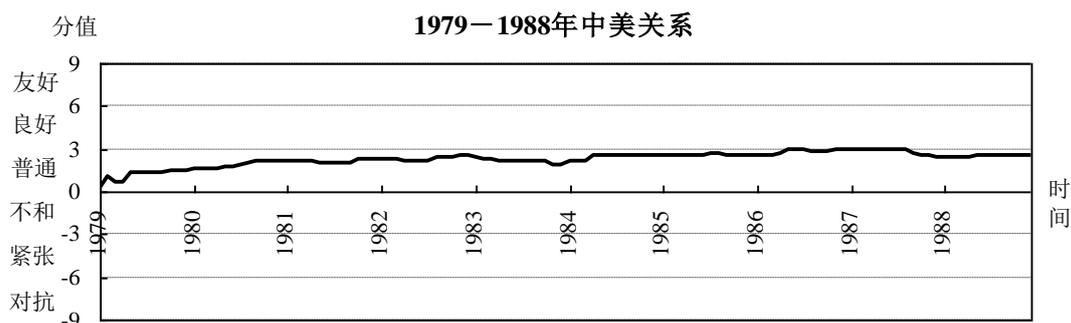


图 1—2

2、比较同一双边关系在不同时期的变化方向。例如，观察图 1—3 中 1994 年和 1995 年的中美关系曲线，读者可看到 1994 年中美双边关系为上升趋势，而 1995 年的双边关系则是下滑趋势。

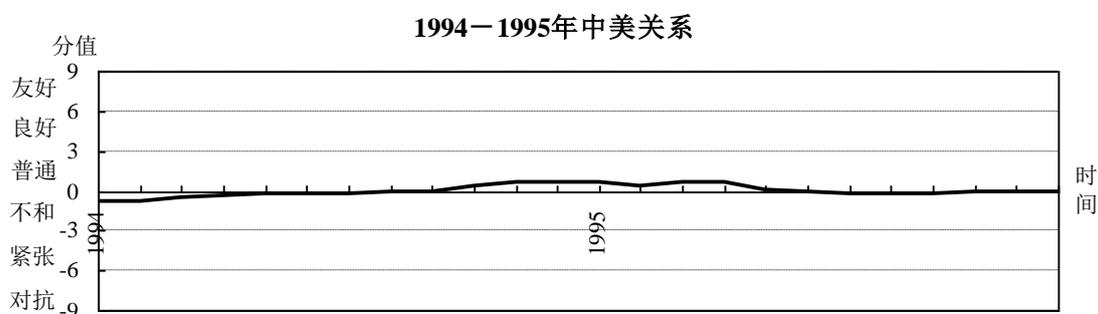


图 1—3

3、比较同一双边关系在不同时期的稳定程度。例如，观察图 1—4 中 1980—1988 年和 1990—2005 年的中美关系曲线，读者可发现中美关系在 1980—1988 年振荡较少，而在 1990—1999 年间则振荡很多。这说明中美关系在 1980—1988 年期间比 1990—1999 年稳定。

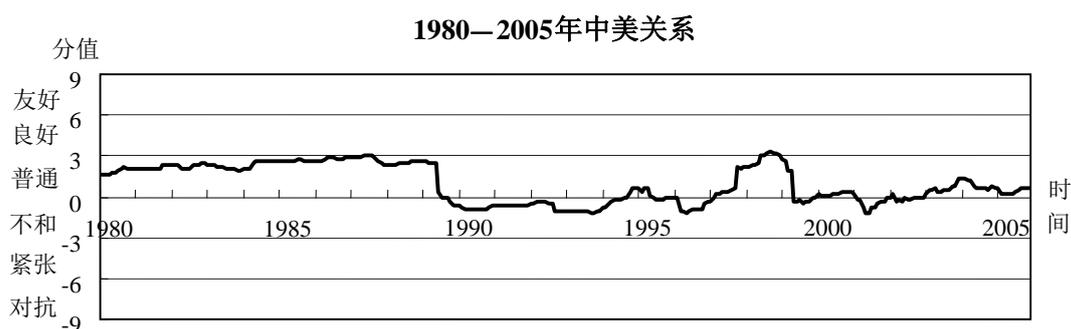


图 1—4

4、比较同一双边关系在同长时间内的变化程度。例如，观察图 1—5 中 1971—1973 年和 1978—1980 年两个时期的中美关系曲线，读者可以发现曲线在 1971—1973 年间是从对抗等级的中等水平的-7.1 升至紧张等级的低等水平的-3.9，升了一个等级多（升幅为 3.2）；在 1978—1980 年间曲线从不和等级的高等水平的-2.6 升至普通等级的高等水平的 2.2，上

升了一个半等级还多（升幅为 4.8）。这说明 1972 年后双边关系改善的速度小于 1978 年中
美建立大使级外交关系后的上升速度。

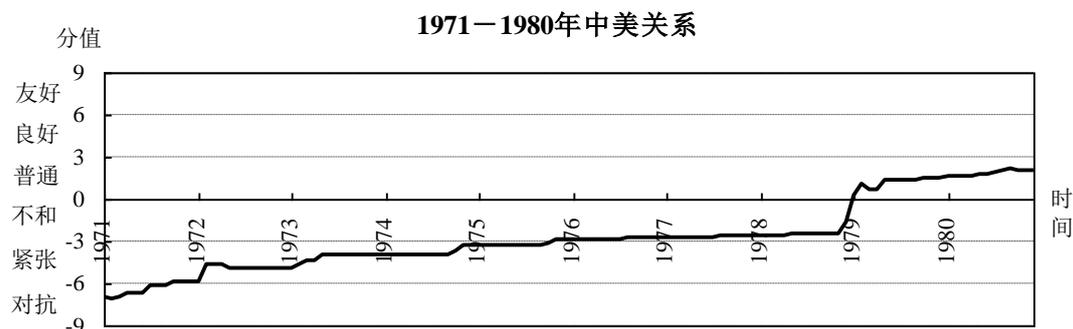


图 1—5

二、两个或多个双边关系比较

两个或多个双边关系比较是对不同双边关系曲线在相同时期的位置、方向、振频和变化幅度进行的比较。和单一双边关系比较一样，也是通过观察曲线位置比较关系的好坏，通过观察曲线方向比较关系走势，通过观察曲线振频比较关系的稳定程度，通过观察曲线在同长时间内的变化幅度比较关系的变化速度。通过观察中国与美、日、英、法四个西方大国的关系曲线，我们可做上述四个方面的比较。

1、比较不同双边关系在同一时期的水平高低。观察图 1—6 中 2005 年的四条双边关系曲线，读者可以看到中法关系最好，位于友好等级；中英关系次之，位于良好等级；中美关系再次，位于普通等级；中日关系最差，位于不和等级。

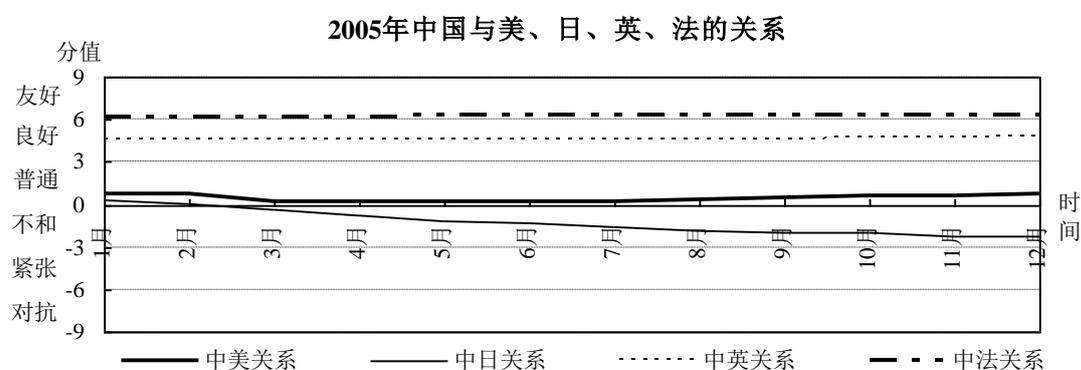


图 1—6

2、比较不同双边关系在同一时期的走向。观察图 1—7 中的四条双边关系曲线，读者可以看到，在 2000—2005 年期间，中法关系与中英关系呈改善趋势，中美关系时好时坏，而中日关系则为恶化趋势。

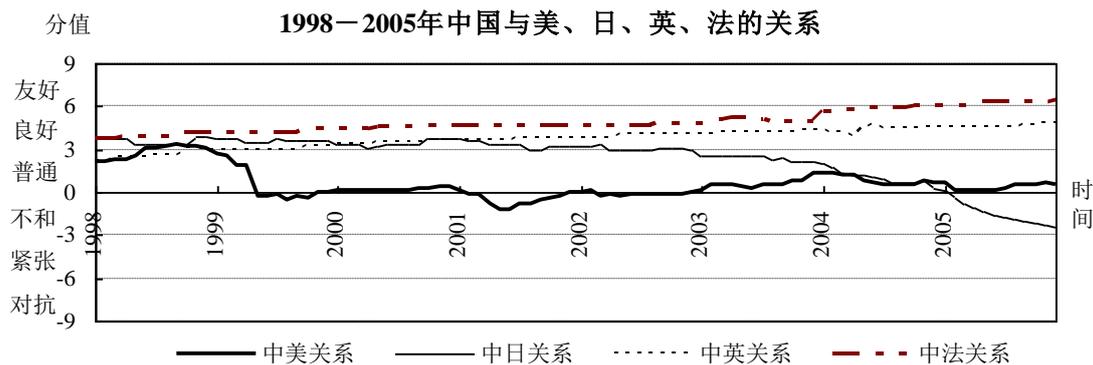


图 1—7

3、比较不同双边关系在同一时期的稳定性。仍观察上图中的曲线，读者可以发现，中法关系和中英关系的两条曲线的振频都很低，中美关系曲线和中日关系曲线的振频很高。这表明，在 1998 年后的这四对双边关系中，中法关系和中英关系比中美关系和中日关系稳定。

4、比较不同双边关系在同一时期的变化速度。观察图 1—8 中的四条曲线，读者可以看到，从 1989 年的年初到年底，中美关系曲线下降最多，下降了 3.3；其次是中法关系曲线，下降了 2.1；再次是中英关系曲线，下降了 1.5；中日关系曲线下降最少，只降了 0.7。这说明政治风波之后，中美关系受损最严重，中法关系次之，中英再次之，中日关系受到的影响最小。

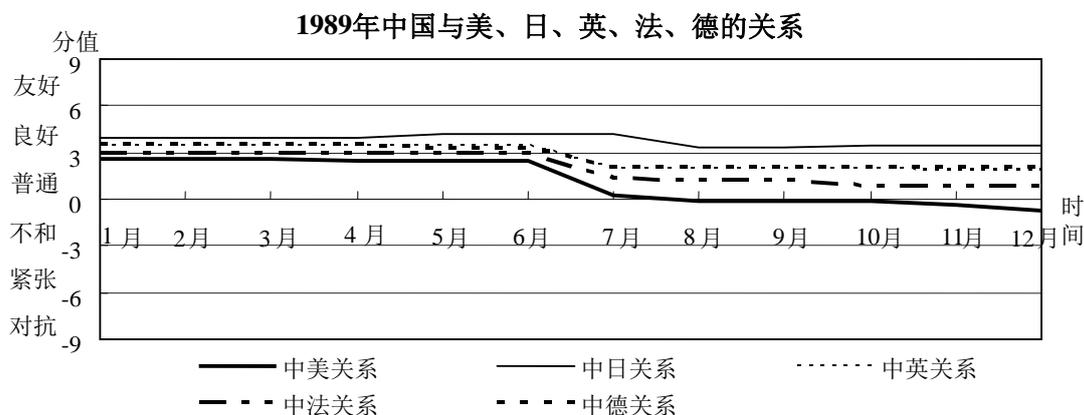


图 1—8

除了使用本书提供的各种图形外，读者还可以根据研究需要，选取本书的双边关系分值制作各种曲线图，以便进行多种不同的观察和比较。所选数据的时间跨度完全根据研究者的需要来决定。

如何使用分值表

一、分值比较

分值比较是比较各双边关系在各时期的分值。例如，查阅表 3—1：1950—2005 年中美关系分值，读者可以发现，1989 年 5 月中美关系分值为 2.6，1989 年 6 月中美关系分值为 0.4。由此，读者可知中美关系在 1989 年 5 月的水平比 6 月高 2.2，双边关系一个月之内下降了两个水平还多。（见下表）

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
1989	2.7	2.7	2.6	2.6	2.6	0.4	0	0	0	-0.4	-0.7	-0.6

再查阅表 4—1：1950—2005 年中日关系分值，读者可以看到，1989 年 5 月中日关系分值为 4.3。由此可知 1989 年 5 月时，中日关系好于中美关系，其绝对差距为 1.7，相差半个等级还略多。（见下表）

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
1989	4.1	4.1	4.1	4.3	4.3	4.3	3.4	3.4	3.4	3.4	3.4	3.4

二、分值等级比较

分值等级比较是比较不同分值所对应的双边关系等级程度。例如，通过查阅表 3—1 中美关系分值，读者已知 1989 年 5 月中美关系分值为 2.6，达到普通等级的高等水平；1989 年 10 月为 -0.4，处于不和等级的低等水平。于是读者可知，中美关系五个月内发生了等级变化。再查找表 4—1：1950—2005 年中日关系分值，读者可发现 1989 年 10 月中日关系分值为 3.4。由此可知，1989 年 10 月，中日关系和中美关系分别处于良好与不和两个不同的等级，其绝对差距为 3.8，相差一个等级还多。

三、变化速度比较

变化速度比较是计算同长时间内的双边关系分值差距。例如，通过查阅表 3—1：中美关系分值可知，1989 年 5—6 月，中美关系由普通等级的高等水平的 2.6 降到普通等级的低等水平的 0.4，变化幅度达 2.2。1999 年 4—5 月，中美关系由普通等级的中等水平的 1.9 降到不和等级的低等水平的 -0.3，降幅达 2.2。由此可知 1999 年 4—5 月中美关系变化速度与 1989 年 5—6 月相同。（见下表）

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
1989	2.7	2.7	2.6	2.6	2.6	0.4	0	0	0	-0.4	-0.7	-0.6
1999	2.7	2.6	1.9	1.9	-0.3	-0.3	-0.1	-0.4	-0.2	-0.3	0.1	0.1

定量衡量的等级概念

定量衡量双边关系的方法是以分值表达政治含义，其目的是加深我们对双边关系政治性质的理解和认识。为此，我们需要明确双边关系程度变化与性质变化之间的关系，也就是说，需要明确双边敌友关系发生何种程度变化会导致关系的性质发生变化。根据敌友性质，国家双边关系可分为“敌对”、“非敌非友”和“友善”三类。我们将每类关系均分为两个等级，这样就得到六个等级不同的双边关系。敌对关系分为“对抗”和“紧张”，非敌非友关系分为“不和”和“普通”，友善关系分为“良好”和“友好”。（见图 1—1）

在定量衡量双边关系时，我们需要严格定义“对抗”、“紧张”、“不和”、“普通”、“良好”、“友好”这六个词汇的等级含义。然而有两个因素会使读者对这六个等级的含义产生不同的理解。第一个因素是以单一事件判断国家的总体双边关系。双边关系是综合性的，包括了政治、经济、安全、文化等多方面。一个方面发生的事件并不能反映双边关系的全部。例如，人们看到两国发生军事冲突时，就认为两国关系已是全面对抗。事实上，在双边关系紧张而未达到对抗状态时，两国也有发生军事冲突的可能。例如，1993 年韩国开始实行和平统一政策，朝韩关系从对抗变为紧张，2000 年朝韩实现了首脑会晤，2002 年双方代表队还在亚运会上同举一面旗帜入场。¹然而在这 10 年间，双方在海洋边境线上却发生过多次军事冲突，但这些军事冲突并未使朝韩关系重新回到全面对抗的状态。此外，没有发生军事冲突的国家也不必然就不处于对抗状态。朝鲜战争结束于 1953 年 7 月，1953 年下半年和 1954 年初中美之间没有发生大规模军事冲突，但此时中美关系仍处于全面对抗的状态。

第二个因素是生活语言的不严格性。在日常生活中，人们使用这六个词汇时是非常随意的和松散的。人们经常用不同的词汇描述相同的关系，或是用相同的词汇描述不同的关系。2004 年，当中国政府公开表态反对日、德、印、巴四国集团的联合国修改方案后，媒体对中日关系的评论各式各样，有的说“中日不和公开化”，有的说“中日关系紧张”，有的说“中日开始对抗”。由于没有严格的定义，同一时期的中日不友好关系被描述成了三种不同的状态。相反，人们用“紧张”一词描述 2003 年法国反对美国布什政府发动伊拉克战争时的美法关系，2005 年美国批评俄罗斯政府脱离民主政治时的美俄关系，2006 年委内瑞拉政府威胁要驱逐美国大使时的委美关系。而事实上这三对双边关系有着本质上的不同，美法在北约内是军事盟友，美俄在八国集团内是战略对话伙伴，委内瑞拉则被美国视为和古巴一样的敌手。

为了有助于读者使用这个手册，我们在此对这六个等级的概念进行明确定义，并就双边关系判断中的常见误解作重点解释。

一、对抗

对抗是指两国关系的性质是敌对的，而且公开称对方为自己的战略敌人。1950 年中美在朝鲜战争时的双边关系是敌对性质。人们常忽视军事冲突规模对国家关系性质的影响。两国处于战争状态，但两国的对抗水平却可能有高低区别。例如，中美在上世纪 50 年代的朝鲜战争中的对抗程度，就远大于 1979 年中越边境自卫反击战时期的中越对抗程度。当时中越虽然也处于战争之中，但是中越保持着正式外交关系，而朝鲜战争时中美没有外交关系。双方发生全面战争时，两国敌对关系的性质一定是对抗的，但两国发生小规模军事冲突，特别是边界军事摩擦时，两国敌对关系的程度未必达到对抗的程度。因为，两国发生军事摩擦

¹ 《世界知识年鉴 1993/1994》，世界知识出版社 1994 年版，第 85 页；《世界知识年鉴 2001/2002》，世界知识出版社 2001 年版，第 98 页。

时，他们可能在非军事方面还进行着合作。例如 2006 年 7 月，朝韩边界发生交火，但朝韩关系并未因此降至对抗等级。²

二、紧张

紧张是指两国关系的性质是敌对的，但没有直接的大规模军事冲突。1972 年中美进行了首脑会晤，但中国还在支持越南的抗美战争，此时的中美关系就处于紧张状态。人们常常误将军事摩擦作为关系紧张的标志，这是不符合实际情况的。双方之间有小规模军事摩擦是紧张状态，没有摩擦时双方不断以战争或军事打击相威胁也是紧张状态。紧张状态分为不同水平，在紧张状态严重的情况下，双边有军事冲突危险，但在不严重的情况下，则可能没有军事冲突危险。官方严厉谴责对方是双边关系恶化的标志，但这种公开谴责的次数和语言内容却决定了激烈程度上的区别。一般讲来，在短期内，全面而连续地严厉谴责对方，两国关系才会达到紧张等级。那些局限于个别问题的严厉谴责并不表明双边关系已达到紧张等级。例如，自 1989 年起，中国政府每年都发表声明谴责美国干涉中国的人权政策，但中美总体关系却是在不和与普通等级之间上下起伏，并未达到紧张的等级。1998 年中美讨论建立“建设性战略伙伴关系”时，中国也未停止谴责美国的人权政策。³

三、不和

不和是指两国关系中的敌对成分大于友好成分，在较多方面采取敌对政策。2001 年中美虽然在经济和防扩散方面有战略合作，但军机相撞使双方军事对立加剧，此时中美关系处于不和状态。人们常将有正式外交关系误解为是友大于敌的标准，其实两国之间是否有外交关系并不能作为国家关系性质或等级的判断标准，因为国家不仅在友好、良好和普通等级关系时保持正式外交关系，在对抗、紧张和不和的关系中也可能保持外交关系。例如，1931 年“九·一八”事变后中日形成对抗关系、冷战后期美苏是紧张关系、2001 年军机相撞后中美是不和关系，但在这些不同等级的关系中，双方都保持了正式的外交关系。因此，双方拥有正式外交关系并不证明双边关系一定是友好大于敌对。以 2005 年不和的中日关系为例，双方经济关系不断发展，贸易额比上年增长了 9.9%；⁴双方文化交流继续扩大，人员往来仍在增多；但是双方在历史问题、海洋权益问题、联合国改革等问题上的矛盾却是敌对性质的。在两国关系中，安全关系和政治关系属于高级政治，经济关系和文化关系属于低级政治，因此 2005 年的中日关系处于敌大于友的不和等级。

四、普通

普通是指两国关系中的友好成分大于敌对成分，在较多方面采取合作政策。2002 年美国在防范中国崛起的同时与中国进行全球范围的反恐合作，此时的中美关系就是普通等级。对于普通关系的误解恰恰与不和关系相反。我们在外交实践中经常使用“关系正常化”一词。这一词语会使人们以为两国关系正常化了就达到了普通等级；两国建立了正式外交关系，就说明两国敌对成分小于友好成分了。然而，事实并非如此。两国实现关系正常化或建立外交关系只表明两国关系的改善，而不能说明两国关系的水平，因为这种改善可在关系不和的情况下完成。也就是说，在双方关系正常化或是建立外交关系时，两国关系在许多方面仍可能具有敌对性质。例如，中美在 1972 年实现关系正常化时，双方在越南战场上还在进行对抗，直到 1978 年才建立大使级外交关系。⁵中日在 1972 年建交时，双方未能就二战和约问题达

² 华亭：《朝韩交火，双方打出 9 发子弹》，《环球时报》，2006 年 8 月 2 日，第三版。

³ 中华人民共和国外交部政策研究室编：《中国外交 1999 年》，世界知识出版社 2000 年版，第 443—449 页。

⁴ 《中国进出口商品主要国别（地区）统计》，《国际贸易》，2006 年 2 月号，第 71 页。

⁵ 韩念龙主编：《当代中国外交》，中国社会科学出版社 1987 年版，第 222—229 页。

成协议，拖延至 1978 年才签署《中日友好和平条约》。⁶因此，正式外交关系的有无，不能作为双边关系是否达到普通等级的标志。

五、良好

良好是指两国关系的性质是友善的，双方满意于现行关系，但在一些问题上有明显的战略分歧。1987 年为了共同对抗苏联的军事威胁，美国帮助中国改进军事装备，此时中美关系就是良好等级。人们常常误以为关系良好的国家没有战略分歧，然而双方满意于现有的关系并不等于双方没有战略分歧。两国没有战略分歧是极少见的现象，即使是军事盟友也会有战略分歧。例如，美国和德国都是北约的成员，但在 2003 年德国坚决反对美国发动对伊战争，美国在 2005 年的安理会改革问题上坚决反对德国成为安理会常任理事国。双边关系是否良好不在于两国是否有战略分歧，而是取决于战略分歧是否足以改变双边关系的友好性质。2005 年和 2006 年中国分别与法国和英国相互举行了对方的文化年，中法和中英关系都处于良好等级，但是同期在伊朗核问题上中国与法、英却持不同的立场，中国不同意对伊朗进行经济制裁。⁷2005 年，中印双边贸易额比 2004 年增长了 37.4%，⁸双方都满意于当时的双边关系，并决定将 2006 年定为中印友好年；⁹与此同时，中印在边界问题和联合国改革问题上仍有重大分歧。

六、友好

友好是指两国关系的性质是朋友，双方国际战略立场基本一致，分歧较小。前面我们以不同时期的中美关系举例说明敌友关系的等级区别。然而，自 1950 年中国共产党取得政权至今，中美关系从没达到过友好等级。所以我们只能用中苏关系为例。1950 年中苏共同支持朝鲜的反美战争，此时的中苏关系是友好等级。人们会误以为友好国家之间没有利益分歧，这是不符合客观事实的。国家间的友好关系是以两国共同利益为基础，但两国利益不可能完全一致，两个国家利益完全一致就会合并为一个国家了。20 世纪 50 年代，中苏关系是友好等级，但双方仍在联合潜艇舰队和长波电台等军事主权问题上存在分歧。¹⁰由于国家之间存在战略分歧是一个普遍现象，友好的国家关系只表明两国能在多数重大战略问题上采取共同立场和行动，两国间的战略分歧没有动摇两国战略合作的危险。

双边关系衡量方法

一、衡量双边关系的基本流程

衡量可以被定义为依据一定的规则，把数字赋予对象和过程。这与人们对对象和过程进行命名在含义上具有相似性。¹¹对于研究变量的测量，当衡量规则确定以后，由于自变量通过一定的规则有惟一的一个量与之对应，这样，衡量的结果也随之确定。¹²

要有效衡量国家双边关系友好或敌对程度，其核心问题是解决从事件赋值向关系赋值的转化。我们的基本设想是，双边关系是由众多事件组成的，这些事件随着时间的推移形

⁶ 同上，第 294—296 页。

⁷ Aliakbar Dareini, “Iran Rejects EU Bid to Halt Enrichment”, http://news.yahoo.com/s/ap/20060514/ap_on_re_mi_ea/iran_nuclear_82.

⁸ 《中国进出口商品主要国别（地区）统计》，《国际贸易》2006 年 2 月号，第 71 页。

⁹ 《孙玉玺大使出席中印关系研讨会》，中国驻印度大使馆供稿，2005 年 4 月 30 日，<http://www.chinaembassy.org.in/chn/sgxw/t222223.htm>。

¹⁰ 韩念龙主编：《当代中国外交》，第 112—113 页。

¹¹ Sandra W. Pyke and Neil Mck. Agnew, *The Science Game* (Fifth Edition) (Englewood Cliffs, NJ: Princeton Hall, 1991), p. 171.

¹² 范克新编著：《社会学定量方法》，南京大学出版社 2004 年版，第 3 章。

成了一个“事件流”，对双边关系的衡量需要在事件累积和流动两个维度上同时进行衡量。也就是说，对事件影响力进行累积是我们衡量的起点，测量事件影响力随时间流动的变化是衡量的过程，双边关系现状是衡量的终点。将事件分值转化为双边关系分值的合理性，将取决于转化后的双边关系分值能否与实际经验相一致，至少要有很强的相似性。为此，我们设计了如下定量衡量双边关系的流程。

定量衡量双边关系的流程

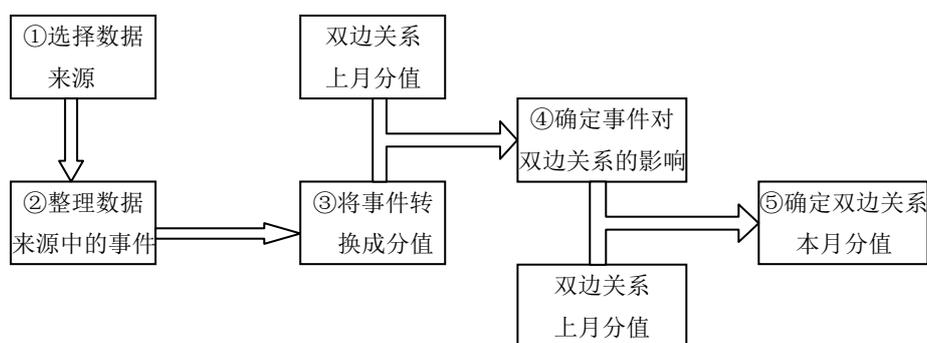


图 10—3

在图 9—3 中，环节①是对数据来源进行选择。从统计分析的角度来说，无须穷尽所有的公开数据来源，只要是数据来源可靠就行。例如，衡量中国与美国的双边关系，选择《人民日报》和中国外交部网站的事件数据就可基本满足需要。

环节②是将数据来源中与我们的研究对象和研究范围有关的事件进行筛选和归类。

环节③是在“冲突—合作”这一维度上，根据对双边关系的影响力大小将已发生事件转换成一维的事件分值。这一转换需按预先确定的事件分值表来实现。设定事件分值表是为了保证转换过程具有较大的一致性和可比性。

环节④是把事件分值进一步转换成对双边关系的影响值。环节④与环节③的区别在于，环节④的赋值需要考虑双边关系所处位置的影响。这是因为，当双边关系的好坏程度处于不同水平时，同一事件对该双边关系的影响力大小不同。例如，2001 年中美军机相撞事件，对两国关系产生了很大的负面影响，但是在朝鲜战争期间，中美间的一场空战就是平常事件，对两国关系的影响极小。我们采取的实际做法是，对于正面事件和负面事件都先用事件影响力公式进行转换，然后再加总得到当月的事件总分值，在此基础上得到该月事件对双边关系的影响值，具体的过程将在后面叙述。

环节⑤是把由事件导致的双边关系的变化值叠加在上个月双边关系的分值上，得到当月的双边关系分值。在这里，我们采取的实际做法，并不是事件影响力分值 0.1，就让双边关系同向变化 0.1，而是事件影响力分值每达到 0.5 让双边关系同向变化 0.1，类似地，事件影响力分值达到 1 则双边关系同向变化 0.2。¹³

二、对双边关系分值标准的设定

用分值表示双边关系的好坏程度的第一个问题是如何制定分值衡量标准，也就是要设定两国关系变化的度量范围。对于数学上的距离空间，人们可以设定许多不同的距离标准，如欧几里德距离，闵可夫斯基距离等等。设定此类衡量标准，要符合方便性和有用性原则。

¹³ 对于这一做法的含义，本节第 6 小点还将进行说明。另外值得注意的是，事件赋值范围的-9 到 9，与双边关系变化范围的-9 到 9，两者之间并没有对应的关系，只是偶然的巧合。

例如在常压下，我们将水的冰点设为摄氏 0°C，沸点设为摄氏 100°C。借鉴这个方法，我们设定衡量双边关系变化的分值范围为-9 到 9，其中，-9 代表两国关系最恶劣的情形，9 代表两国关系最友好的情形，这是两种极限情况。两者的中值为 0，它表示两国关系处于绝对的非敌非友状态，两个毫无关系的国家之间的关系可视为这种状态（如柬埔寨与布基纳法索的关系），或是双边关系中的合作与冲突绝对地各占 50%的情况。¹⁴关于国家双边关系分值标准见本书第 1 部分“使用方法”。

三、事件分值表的设定

为了避免赋值的随意性，我们依照事件数据分析中的通行做法，设定一个确定的事件分值表来保证赋值的一致性。¹⁵在事件数据分析中，比较具有代表性的分值表有阿萨尔-斯隆分值标准、戈尔德施泰因分值标准和“世界事件互动测量”分值标准。但这些分值标准主要是用于分析冲突行为的，无法用于衡量双边关系，为此我们制定了一套与我们设定的双边关系分值标准相一致的事件分值标准（参见本书第 10 部分“事件赋值标准”）。¹⁶

设定这个事件分值表时，我们坚持了完整性和排他性两个原则。完整性是指可能选取的事件都能按分值表赋予分值；排他性是指同一事件只适用于一个分值，避免同一事件归入两个以上类别的可能。事件分值表是根据经验确定的，但是在它确定以后不再改变，从而避免赋值的主观性和随意性。

四、事件影响力分值的计算

事件对双边关系分值的影响力取决于本月两国关系中的事件和两国关系上月的分值这两个变量。事件的作用是使两国关系发生变化（变好或变坏），但这种变化是在现有两国关系基础上发生的。在不同水平关系的基础上，相同事件对两国关系的影响力不同。例如在两国关系已经很友好的情况下，首脑会晤对双边关系的改善作用是很有限制的。而在两国关系不好的情况下，举行一次首脑会晤，或者签订一项合作条约能起到明显改善双边关系的作用。为了解决事件在不同关系水平上的影响力不同的问题，我们根据双边关系分值标准，设计了如下的事件影响力计算公式。

$$I = \begin{cases} \frac{N - p_0}{N} I_0 & \text{当 } I_0 \geq 0 \\ \frac{N + p_0}{N} I_0 & \text{当 } I_0 < 0 \end{cases}$$

I 表示事件在两国关系位于 P_0 时的影响力分值

N 表示两国关系变化范围的绝对值

P_0 表示事件发生时两国关系的初始值¹⁷

I_0 表示事件在事件分值表中的分值

两国关系的变化范围是 $[-N, N]$ ，在本研究中即为 $[-9, 9]$ 。

上面的公式具有如下特点：1、当事件对两国关系起正向推动作用时，即当 $I_0 > 0$ 时，随着初始位置从对抗向友好的方向移动，事件的正向推动作用逐渐减小。当两国关系达到

¹⁴ 全球化可能使两个国家完全没有关系变得不可能，但有些国家之间的确没有实质性的安全、政治或经济关系。

¹⁵ 事实上，现在的事件数据分析，主要采用机器编码方法，而基本上放弃了手工编码，因此，对编码系统的依赖更强了。

¹⁶ 我们制定的衡量中外双边关系的分值表还不完善，但主要是技术细节问题，并不影响主旨分析，特别是不影响分析的方向和性质。

¹⁷ 由于这里的分析是以月为单位，因此，如果我们考虑 5 月份发生的事情的影响，那么，两国关系的初始值 P_0 就是 4 月底时两国关系的分值。

最友好状态时，正面事件的正向推动作用为 0。例如，1971 年尼克松总统访华之前，中美在越南战场上是敌对方，美国总统尼克松访华使中美关系发生了质的变化，而 1978 年中美建立了正式外交关系后，美国总统里根 1984 年访华对双边关系的推动作用就远远小于尼克松访华的作用。

2、当事件对两国关系起负向推动作用时，即当 $I_0 < 0$ 时，随着两国关系初始位置从对抗向友好方向移动，事件的负向推动作用逐渐增强。在两国关系最紧张情况下，负面事件的负向推动作用为 0，即任何负面事件都不能使两国关系的分值小于 $-N$ 。例如，2000 年美国小布什上台，此时中美关系是非敌非友，因此撞机事件使得双边关系严重倒退，但在 1965-1971 年援越战争时，双方互为敌手，中国军队在越南战场上打下美国轰炸机或是美军轰炸中国防空阵地，对双边关系都没有重大影响。

3、该公式具有对称性。当 P_0 与 I 的分值相同且符号相反时，意味着影响力相同，但方向相反。

五、两国关系当前分值的确定

两国关系的本月分值等于两国关系的上月分值加上由事件引起的本月两国关系的变化值。这似乎是一个简单的相加，但由此会导致一个理论上的问题，即上月分值如何确定？因为上月分值的确定又依赖于上上个月的分值，这意味着有了上上个月的分值才能获得上个月的分值。这一逆推过程可以无休止的，于是产生了第一次主观确定的初始值误差影响以后分值计算准确性的问题。然而，我们所设计的影响力公式具有一种纠偏能力，随着统计时间的延长可以纠正初始值误差，从而解决了确定上个月分值的准确度难题。下一节我们将通过检验双边关系初始值收敛效果的方法来表明这一点（参见图 9-4）。

六、具体操作说明

这一部分对具体的衡量过程做一点更详细的说明，从而使读者能够直接运用本书中的数据计算双边关系在各个时段的分值。这里的步骤是指已经对事件进行了整理以后，如何根据这些事件计算关系分值的步骤。

1、计算当月事件原始分值

以 2004 年 4 月的中美关系为例，根据我们收集的事件，并按照事件分值表赋予分值，得到该月中美关系的事件分值分别为正向事件分值 2.6，负向事件分值 -1.9。

2、对事件原始分值进行转换

我们在计算当月事件分值的时候，没有采用把正向事件和负向事件的分值直接相加的方法。其原因在于：当双边关系处于不同位置时，正向事件和负向事件的影响力是不同的。例如，自 20 世纪 90 年代以来，中俄关系一直处于较高的水平，此时，如果中俄关系中某个月的正向事件分值和负向事件分值分别为 2 和 -2，并不意味着该月的事件分值应该为 0，实际上，这意味着该月双边关系的势头不好。即此时负面事件对双边关系的负面影响大于正面事件对双边关系的正向影响。

基于这一考虑，我们采取的实际做法是，根据当时双边关系所处的水平，对事件分值按照前述事件影响力公式进行计算，然后把计算出的分值相加，得到当月的事件分值。由于 2004 年 3 月中美关系的分值水平为 1.2，按照事件影响力公式，2004 年 4 月中美关系中正面事件分值总合为 2.6，其正向影响力分值为 $\frac{N - P_0}{N} \times I_0 = \frac{9 - 1.2}{9} \times 2.6 = 2.3$ ，该月负面事件的分值总合为 -1.9，其对双边关系的负向影响力分值为 $\frac{N + p_0}{N} \times I_0 = \frac{9 + 1.2}{9} \times (-1.9) = -2.2$ 。两者合并得到该月事件总影响力分值为 0.1。

3、把事件影响力分值的变化折算成双边关系分值的变化。

如前所述，在我们的设计中，事件影响力分值 0.1，并不意味着双边关系同向变化 0.1，我们制定的规则是，事件影响力分值每增加 0.5 时双边关系同向变化 0.1。按照这一规则，2004 年 4 月事件的 0.1 影响力分值折算后只有 0.02。不足 0.1 的最小计算单位省略不计，因此 2004 年 4 月中美关系变化分值为 0，与上月分值相加后得到 4 月份双边关系分值仍为 1.2。¹⁸

事件影响力达到 0.5 时才让双边关系变化 0.1，其含义是把事件分值表中的事件分值全部缩小了 5 倍。这种折算方法在数学上或逻辑上是等价的，不会产生实质性差异。采用这种折算方式还可以避免双边关系分值超越 9 和 -9 的边界。虽然，随着双边关系越来越接近 9/-9，同向事件的影响力越来越小，反向事件的影响力越来越大；但是，一个月内如果发生多起重大反向事件，双边关系分值则有超过 9/-9 的可能性。我们的折算方法，就避免了双边关系分值越过边界值的可能。下表为依据我们的折算标准计算出来的，处于不同水平的双边关系超越 9/-9 边界值所需要的事件分值：

表 10-1：从不同双边关系水平越过双边关系上下限所需要的事件分值

关系水平	越过 9 所需要的正向事件分值	越过 -9 所需要的负向事件分值
-9	22.5	$-\infty$ (负无穷大)
-8	23.8	-405.0
-7	25.3	-202.5
-6	27.0	-135.0
-5	28.9	-101.3
-4	31.2	-81.0
-3	33.8	-67.5
-2	36.8	-57.9
-1	40.5	-50.6
0	45.0	-45.0
1	50.6	-40.5
2	57.9	-36.8
3	67.5	-33.8
4	81.0	-31.3
5	101.3	-28.9
6	135.0	-27.0
7	202.5	-25.3
8	405.0	-23.8
9	∞ (无穷大)	-22.5

从上表中我们可以发现一个有趣的特征，即双边关系的水平离一个边界越远，超过边界所需要的事件分值越小。但是，国家间关系的实际情况是，在双边关系很好或很坏的情况下，一个月内同时发生多个极其重大反向事件是不可能的。当双边关系处于 0 水平时，使双边关系越过正/负向边界需要的事件分值为 45/-45，而我们的事件分值表中最重大的单项事件的分值也只有 9/-9。随着双边关系水平的提高或下降，要使双边关系越过反向边界实际上是

¹⁸ 在这里我们可以看到，如果采取把正向事件分值与负向事件分值直接相加，然后再用事件影响力公式进行调整的做法，则计算的结果将是中美关系上升 0.1。

越来越难了。如双边关系分值为 7/-7 时，要越过同向边界，一个月内的同向事件分值需要达到 202.5/-202.5（这还要以没有反向事件为前提）。上述情况表明，依据我们的折算标准，9 和 -9 的边界值是非常难以越过的。虽然理论上存在着分值超越边界值的可能性，但这种情况实际上则不会发生。国家关系变化的渐进性更决定了关系分值越界的情况不会发生。

衡量方法的有效性测试

一、对初始值（上月分值）计算方法有效性的检验

如前所述，在衡量两国关系的当前分值的时候，面临的一个难题是初始值无穷逆推。我们解决这个问题的思路是，对过去某个时间的双边关系做多个任意赋值，然后从这些不同初始分值点出发计算当前分值，经连续时间的统计计算，观察它们各自结果之间有多大差别。如果差别依旧，则说明这种计算初始值的方法无法解决初始值不准确的问题，如果汇集于一点，则说明使用这种计算方法可以矫正初始值赋值不准确的问题。为此，我们做了如下检验，对 1980 年 1 月的中美关系做多个显著不同的赋值，让它们从不同的初始值出发，运用同样的事件数据计算中美关系分值的变化，其结果得到如下的演化图。

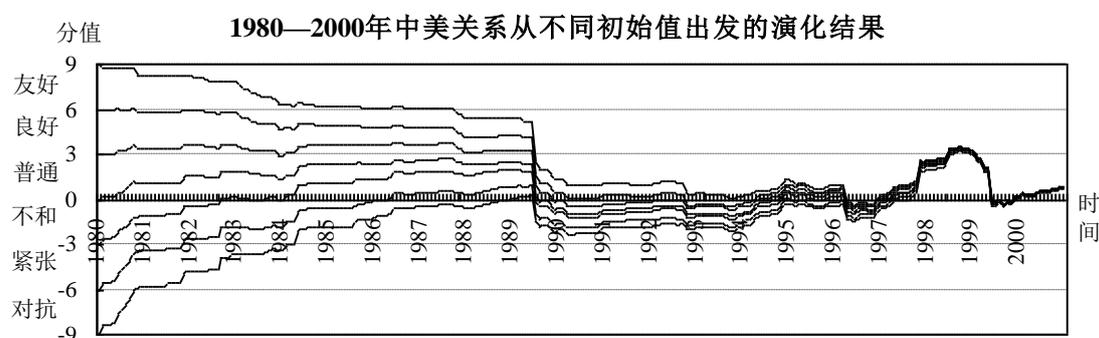


图 10-4

图 9-4 表明，随着时间的延长，从不同分值起点出发的中美关系的每月分值呈现向狭窄范围汇聚的趋势，随着时间的进一步延长，它们汇聚到同一个点。另外，运用数学方法，我们也能证明初始值不影响双边关系分值的未来发展。¹⁹特别是，对于极端的情况（即双边

¹⁹在时间长的前提下，双边关系初始值 P_0 不影响双边关系分值 P 的代数证明如下：

$$\text{根据影响力转换公式： } I = \begin{cases} \frac{N-P_0}{N} I_0 & I_0 \geq 0 \\ \frac{N+P_0}{N} I_0 & I_0 < 0 \end{cases} \Rightarrow I = I_0 - \frac{|I_0|}{N} P_0$$

$$\text{可知 } I_1 = I_{01} - \frac{|I_{01}|}{N} P_0 \quad \therefore P_1 = P_0 + I_1 = P_0 + I_{01} - \frac{|I_{01}|}{N} P_0$$

$$\therefore P_n = P_{n-1} + I_n = P_{n-1} + I_{0n} - \frac{|I_{0n}|}{N} P_{n-1} = I_{0n} + (1 - \frac{|I_{0n}|}{N}) P_{n-1}$$

$$\text{令 } K_n \text{ 为 } P_n \text{ 中含有 } P_0 \text{ 的项，则 } K_n = P_0 \prod_{i=1}^n (1 - \frac{|I_{0i}|}{N})$$

$$\therefore |I_{0i}| \leq N \quad \therefore 0 \leq 1 - \frac{|I_{0i}|}{N} \leq 1 \quad \therefore \lim_{n \rightarrow \infty} \prod_{i=1}^n (1 - \frac{|I_{0i}|}{N}) = 0 \quad (\text{在此必须指出，当 } |I_{0i}| = 0 \text{ 时，即双边关系所有时间的分值恒等于 0，则 } \lim_{n \rightarrow \infty} \prod_{i=1}^n (1 - \frac{|I_{0i}|}{N}) = 1)$$

然而，在现实双边关系中，由于双边关系事

关系从 9 或-9 开始变化), 数据也能够较快地收敛, 这意味着使用本文设计的双边关系衡量方法, 即使对过去两国关系初始值的赋值有误差, 也能将当前双边关系分值水平纠正到与双边关系实际情况相符的点上来。由于这种计算方法对初始值的统计是以以往若干年的“事件流”为基础的, 因此较长时期的连续统计计算结果, 实际上是对“事件流”的浓缩反映, 从而能消除较早初始值的误差, 保证当前初值(上月分值)的准确性。由此也说明这种计算方法是合理的和有效的。客观上, 双边关系初始值的确立是以国际关系的常识为基础的, 因此其差距不会特别大, 可以在比较短的时间内收敛到客观实际情况。由于 1949 年 10 月是中国政权更迭的年头, 因此本书 1950 年 1 月的双边关系分值是由专家讨论后主观给定, 这一分值与事实可能有出入, 但不会很大, 而且我们的计算方法能较快地纠正这一误差。

二、对衡量方法总体效果与实际情况一致性的检验

为了检验这一双边关系衡量方法是否可以客观地反映实际情况, 我们需要用经验事实作依据进行检验。为此, 我们让两个不同的双边关系从相同的初始值出发, 然后进行连续的统计计算, 观察两个双边关系的不同“事件流”是否能够使它们产生差别, 以及这种差别是否与我们所了解的客观历史事实相一致。如果与客观历史事实一致, 则说明这套双边关系定量衡量方法是有效的, 否则是无效的。

我们以中美关系和中俄关系为例进行检验。为不失一般性, 我们设这两个双边关系的 1980 年 1 月的初始点分值都为 0。然后, 根据前述衡量流程对这两个双边关系的分值进行逐月统计计算, 得到如下演化图。

件的影响, 双边关系所有时间的分值不可能恒等于 0, 所以 $\lim_{n \rightarrow \infty} \prod_{i=1}^n (1 - \frac{|I_{0i}|}{N}) = 0$ 成立。))

故知 $\lim_{n \rightarrow \infty} K_n = 0$, 即当时间越长 (n 越大) 时, 双边关系初始值对双边关系分值的影响越小, 直至不产生影响。

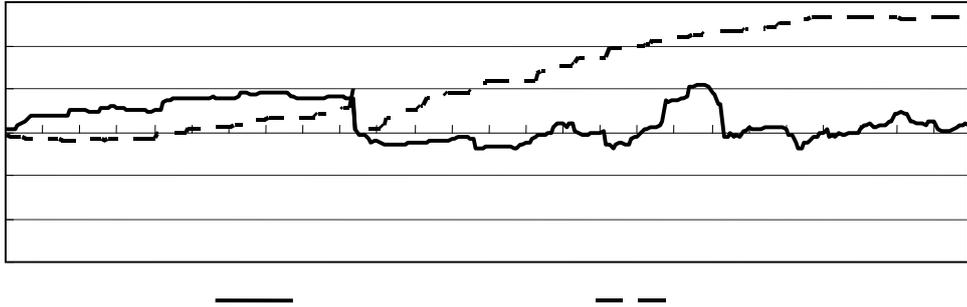


图 10—5

从图 9—5 中我们可以看到，1980 年 1 月中美关系和中俄（苏）关系从相同的初始值出发，但它们的演进路线很快就分岔了。历史经验是，在 20 世纪 80 年代，中美关系明显好于中苏关系，中苏关系到 1985 年后出了改善的趋势。在 90 年代，中俄关系得到了长足的

发展，而中美关系则处于较低的水平上，且摩擦较多。图 9—5 显示，同一初始值出发的两条曲线在发展趋势上的差异基本上与我们对这一时期中俄（苏）关系和中美关系的历史经验相一致。中俄关系曲线的平滑和中美关系的曲折，也符合这一时段中俄（苏）关系与中美关系的实际历史特征。这说明用“事件流”定量衡量双边关系是可以反映客观实际情况的。

从图 9—4 和图 9—5 所显示的检验结果中，我们还可以看到，初始值的位置离当前值位置越远，则当前分值的误差越小。这表明，用这种衡量方法得到的双边关系当前值，能较多地包括双边关系中的新近事件信息，能较充分地反映当前双边关系的实际状态。这对衡量当前双边关系来讲，是一种很好的性质。

三、对双边关系轻微变化的敏感性检验

这一衡量方法对双边关系的敏感性高于人的主观感觉。有时候，双边关系的变化是轻微的，人们主观上不易察觉到这些变化。当双边关系的这种轻微变化向同一个方向积累了较长时间后，人们才能意识到双边关系的变化方向。如果使用我们设计的衡量方法，能在人们意识到双边关系的变化方向之前描绘出双边关系的变化方向，则表明这种方法对双边关系变化的敏感性强于人们的主观感觉。

我们近几年对中日关系变化的衡量是一个有力的例证。下图是根据本文所述的衡量方法得到的 2000—2005 年中日关系的走势：

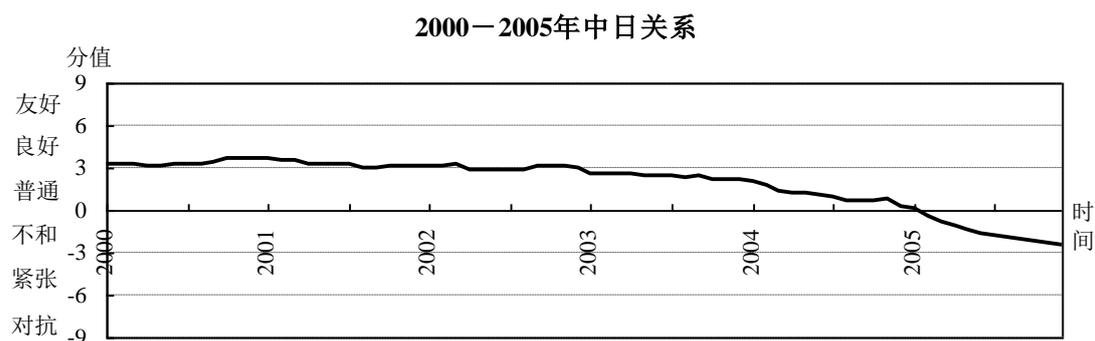


图 10—6

从图 9—6 中可以看出，在 2003 年，中日关系虽然总体下降，但中间还有若干小的起伏；2004 年以后，中日关系开始了一路下滑的过程。我国学界和媒体是到了 2005 年才普遍意识到中日关系出现了下滑趋势。而我们在 2004 年就根据双边关系衡量的结果，撰文指出中日关系出现了下滑趋势。²⁰我们的判断在当时还得不到人们的认同，因为当时的中日关系中的一些问题没有像后来那样集中地表现出来，给人们主观感受的冲击不够大。我们当时就公布了衡量结果，并据此判断中日关系的下滑势头。这在当时有一定的风险。²¹事实上，如果我们当时不做定量衡量，也难以在 2004 年较早地意识到中日关系的下滑趋势。这个事例客观性很强，它表明本书设计的定量衡量方法是能够较敏锐地捕捉双边关系微弱变化的。

上述几点表明，用这一衡量方法得到的结果与经验事实基本吻合。这并非偶然和巧合，而是表明这一衡量方法基本上是有用的。

对一些问题的说明

²⁰ 周方银：《从预测误差看中日关系走势》，《世界知识》，2004 年第 22 期，第 46—47 页。

²¹ 从这种意义上说，这一衡量方法是具有经验内容的，是具有可证伪性的。

一、关系分值和对分值的主观感受²²

在上一节，我们讨论了衡量的效度问题，即衡量得到的分值是否与所讨论对象的不同变化相对应，从而能够显示所要衡量对象的属性。²³下面我们将讨论衡量的信度问题，即同样的方法，被重复地应用于对同一对象的测量时，是否每次能够产生同样的结果。²⁴

使用本书设计的衡量方法，所得双边关系分值有较高的信度。衡量是以一种稳定和内在一致的方式对对象赋予分值的过程。如果这种衡量有助于我们区分出双边关系的不同水平，并具有公开性和可重复性，则这种衡量就具备了普遍接受的条件。由于缺少统一标准，因此许多主观因素使人们对同一双边关系事件的重要性产生不同的认识。如果按照一个固定的事件分值标准对事件赋予分值，就可以大大降低主观因素影响，使人们对同一事件重要性的赋值一致。我们在确定事件分值标准时，参照历史经验进行了反复的斟酌。当分值表确定后，我们对于同类事件赋予的分值都是一样的。这在很大程度上保证了衡量结果的可重复性和稳定性。

从当事人或研究者的角度来看，同一类事件的重要性实际上是有差异的，例如，同为国事访问，但每次访问对双边关系的影响实际上是不同的。由于两方面的原因，我们将忽视这种不同：（1）对于这种差异具体有多大，研究者常常难以清晰地表达出来，甚至两次国事访问相比，哪一次更为重要，人们也常常不能取得共识。（2）我们寻求的是一种具有普适性的方法，这种方法要求在相当程度上超脱就事论事的细节研究，从而使其具有可操作性。因此，对于相同事件不同影响的差异，我们主要依靠统计性质来解决。除特殊情况外（如第一次国家领导人的正式访问），我们并不通过具体地进入事件和文本本身来确定事件的重要性，这有助于保持数据的一致性，避免人言人殊的情况。另外，当统计的数据足够多时，一些具有随机性质的误差可以相互抵消。事件分值标准表中的事件权重是基于 50 多年的历史数据和经验给出的，因此它在某种意义上是一个平均值，是对两者在统计意义上的匹配。

二、数据的手工编码和机器编码

我们在这项研究中采用的是手工编码，其信度即可重复性低于机器编码，但手工编码比机器编辑有更强的合理性。首先，机器编码的可重复性强，但错误编码率也高。例如，《人民日报》发表的九评，机器编码很难把它们的重要性与一般社论区别开来；很多社论不点出其所批判的国家名，机器编码无法识别；有些信息源中的用语上的模糊也可能导致机读错误，如在《美国击败中国》之类的标题下介绍的可能是篮球比赛而不是政治问题。²⁵其次，机读会产生重复编码。由于媒体对同一事件会从几个不同的角度进行报导，因此机器阅读时可能会将持续数天的事件当成不同的事件重复编码。对于有限数据的编码，机器编码达不到手工编码的准确度。在数据量有限，编码工作能由人工完成的情况下，手工编码得到的数据更为有效。我们的手工编码是由国际关系专业人士进行的，而且相互之间还进行了交叉检验。本项研究使用手工编码实际上是提高了数据的质量。我们无法保证每个编码绝对无误，但是其失误率将明显低于机器编码，并在可接受范围内。就双边关系的定量衡量来说，效度是一个更为核心的问题。²⁶在衡量双边关系时，有效性高于效率是我们的原则。

²² 秦亚青对这一衡量方法在这方面的性质进行了有价值的讨论。秦亚青：《国际关系的定量研究与事件分析方法——评〈国家双边关系的定量衡量〉》，《中国社会科学》，2005 年第 1 期，第 135—137 页。

²³ 关于效度问题，可以参考艾尔·巴比（Earl Babbie）：《社会研究方法》（影印本），清华大学出版社 2003 年版，第 143—144 页。

²⁴ 艾尔·巴比：《社会研究方法》，第 140—141 页。

²⁵ 由于机读采用的是关键词搜索的方式，这可能导致即使标题中是“美国队”和“中国队”，也被检索出“美国”、“中国”、以及体育比赛中经常出现的“进攻”、“防御”、“胜利”、“失败”等关键词，并进行错误编码。

²⁶ Russell J. Leng, “Automated and Machine-Assisted Coding of Event Data: The BCOW Approach”,

在原始数据方面，除《人民日报》和外交部网站以外，我们还做了一些补充，这一补充仅限于比较重大的事件，如1971年7月基辛格秘密访华，这是中美关系史上的重要事件，但由于这是秘密访问，《人民日报》没有报导，对此我们需要根据其他来源进行补充，以使中美关系的分值更符合实际情况。这种补充的一个弊端是，会对数据口径的一致性产生一些不利影响。为此，我们把这种补充限制在对很少量的重要事件进行补充的范围内。

三、时间单位问题

我们以月为衡量双边关系分值的时间单位。衡量双边关系的时间单位的选择，涉及变化程度和关系变化过程的平衡问题。天、周、月、季度、年均可作为关系变化的时间单位。时间单位太短（如以天为单位），无法观察双边关系的变化程度，因为一个事件在一天内无法完成；时间单位太长（如以年为单位），则不足以观察国家间双边关系变化的详细过程。经过试验，我们发现，跨越一个日历周的国际事件太多，容易造成同一事件多次赋值。我们也以两个月为时间单位进行统计分析，其结果与以一个月为单位的计算只有细微的差异。于是我们选择了以月作为双边关系衡量的最小时间单位。

结论

根据我们对本书设计的定量衡量双边关系方法的检验和应用，其结果显示该方法有较强的客观性、可重复性和准确性。本书所介绍的将事件分值转化成为关系分值的方法，克服了事件分值等同于国家关系分值所产生的分值与实际关系不符的缺陷。固定的赋值表和分值计算方法使衡量结果有了客观标准。本书所设计的方法仅是提供了把事件数据分值转化为双边关系分值的一个思路，在这一基础上，研究者们可以进一步开展工作，对技术细节进行改进和完善，并可以用不同的数据来源来对这一衡量方法进行检验。我们这项研究工作中的最大缺陷是事件数据来源单一。《人民日报》和中国外交部网站的报导有一定片面性，如果能将双边关系另一国的官方数据也纳入数据来源，则可以增强赋值的准确性。但这显然要几倍地增加衡量工作成本，由于我们在人力和物力方面的限制，未能做到这一点。希望国际关系研究界的同行们与我们共同检验这一衡量方法的合理性和可靠性，从而促进定量衡量双边关系研究的深入发展。